Towards improved traceability in the seafood industry.

Marine Laure Ioana Cusa

June 2022

This thesis is presented to the University of Salford, School of Science, Engineering and Environment, in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

Supervisor: Prof. Stefano Mariani
Co-supervisor: Dr. Ian Goodhead

"So, should we race to see how quickly we can consume the last tuna, swordfish, and grouper?
Or race to see what can be done to protect what remains?
For now, there is still a choice."

— Sylvia A. Earle, *The World Is Blue: How Our Fate and the Ocean's Are One* —

Table of Contents

List of Tables

List of Figures

Acknowledgments

When I first read about the SEATRACES project, I immediately knew I wanted to join the team. A month later, I had moved to Manchester and was embarking on a fantastic journey in the Mariani lab, beginning a PhD project linked to SEATRACES. I want to thank you, Stefano, for not only giving me this opportunity but for trusting me in conducting the work despite my lack of experience using molecular biology tools at the time. You were there for any query I might have and were always able to synthesize my sometimes-messy thoughts, but most importantly, you trusted my judgment and gave me room to expand and create my own path. You encouraged me to join the fantastic field work of some colleagues and always supported my ambitions. I cannot explain how much it all means to me. I have learned so much through your unique mentorship and truly wish we can continue to work together for years to come! I am also grateful to all the members of the SEATRACES project for welcoming me as part of this fascinating team!

Ian, you once said to me “You don’t need to thank me, I am your co-supervisor”. Yes, but you see, I do! For your patience and for the time you’ve invested in me. For the time you’ve spent in the lab, relentlessly explaining and teaching me “how it works”, for offering me cake while drawing awkward yet immensely helpful diagrams, and for brainstorming with me on project goals, plans, uncertainties, and problems… Thank you!

There are many beautiful friends I have met along the way, from university and from the climbing gym, without whom life would not have been the same. I thank you all for helping me keep my sanity and integrity during hard times, and for making the fun times so much more… fun! Thanks so much to Charles and Peter for sharing your knowledge, tips, and wisdom, and for the ongoing help throughout the years and in the lab. Without you both, it would have taken me a lot longer to complete this work! I’d also like to thank you Chris, for being such an inspiring mate, for your insight, and for all those thoughtful discussions of which I hope we can have many more! And Andhika, I’m not sure how to thank you for everything you have taught me. This experience with you working on shark

fisheries in Indonesia was incredible, eye-opening, and captivating. You have been a beautiful friend and collaborator, thank you for your kindness and patience, and for showing me the ways of your country and your culture.

Matt, my love, it was not easy to live apart for three years while undertaking this thesis. Yet you supported me all the way through and as you rightfully said, it also allowed us to explore yet another country that we have now both come to love. I cherish every day I spend with you and cannot wait for the new adventure(s!) awaiting us.

Maman, papa, Pierre, mes héros. Vous êtes les piliers qui font de moi ce que je suis aujourd'hui. Votre passion pour le monde, pour l'aventure, pour l'exploration et la découverte coule dans mes veines. Votre amour inconditionnel et votre soutien envers chacun de mes rêves m'a permis de vivre et d'avancer avec certitude et enthousiasme. Vous êtes ma lumière !

Cyril, comment te remercier toi et ton équipe ? C'était une révélation pour moi de bosser avec vous sur le Black Pearl. Vous m'avez ouvert les yeux, sur le monde complexe et fascinant de la pêche et sur le travail incroyable et parfois épuisant des pêcheurs. Vous m'avez plongé tête la première dans cette passion qu'est votre métier et c'est aussi grâce à vous que je suis ici aujourd'hui. Alors merci, merci de m'avoir accepté à bord de votre bébé, de m'avoir fait confiance et de m'avoir tant appris !


BLACK PEARL

General Abstract

Demand for seafood products is increasing worldwide, contributing to ever more complex supply chains and hampering traceability efforts. Despite marked improvements in seafood traceability and transparency, the fisheries industry is still victim to fraudulent behaviours and malpractice. This thesis examines some of the societal factors that may affect seafood traceability and explores DNA-based methods that have the potential to greatly improve the continuous and regular monitoring of transparency and traceability along the supply chain. Each chapter is dedicated to a given driver which might allow mislabelling to persevere (i.e. lack of consumer knowledge, shortcomings of species identification methods, absence of a framework for the use of point-of-origin detection tools) and explores some of the associated solutions that could help strengthen the monitoring of seafood products, verify compliance, and tackle fraud in the seafood industry. Educational tools and DNA-based methods can empower both consumers and enforcement officers respectively in the quest to combat fraudulent practices in the seafood industry; yet most enforcement bodies still struggle to identify which tools to work with. This highlights a potential mismatch between what the scientific community proposes and what the users really need. This thesis attempts to bridge the increasing demand for simple traceability and transparency tools with some of the existing technologies and proposes frameworks and strategies for their adoptions in practical contexts. It emphasises that if interested parties invest in coordinated efforts to develop robust and comprehensive authentication methods for an increasing number of commercial species, the benefits would largely outweigh the costs. Marine resources are under tremendous pressure and the need for good stewardship is now critical. Effective tools do exist, and it is crucial to demonstrate their practical application and expose the reach they may have within the fisheries and seafood industry.

Chapter 1
General Introduction

1.1 Seafood traceability and mislabelling

1.1.1 Fisheries and seafood consumption

The planet's population is expected to reach 9 billion by 2050 urgently demanding that seafood resources are being harvested sustainably. 3.3 billion people worldwide rely on fisheries for their livelihood (FAO, 2020), many for whom the micronutrients found in fish cannot be found elsewhere (Hicks et al., 2019; Thilsted et al., 2016). The seafood demand is imbalanced; some countries with depleted stocks are exploiting and importing seafood from other more resourceful countries whose local communities themselves depend on and are thereby deprived of the consumption of seafood for livelihood (Swartz et al., 2010). Increasing demand for seafood products worldwide (FAO, 2018a, 2018b, 2020) has reached 179 million tons in 2018 and is contributing to ever more complex trade routes and chains of custody, hampering traceability efforts (Cawthorn & Mariani, 2017). This rate of seafood consumption is putting tremendous pressure on marine resources and drives dishonest harvesting. Between 1974 and 2017, the percentage of unsustainably harvested stocks has gone up from 10% to 34% while the percentage of underfished stocks had gone down from 39% to 6% (FAO, 2020). Sustainable stock management depends on compliance to regulations along the supply chain and fraudulent behaviors leading to mislabelling can have a dramatic impact on natural resources, consumers, and on Small and Medium Enterprises (SMEs) who depend on these resources. Although improved European Union legislations (Regulation No 104/2000/EC and No 2065/2001/EC), consumers' awareness, and industry's accountability have contributed to decreasing seafood mislabelling in Western Europe (Mariani et al., 2014, 2015), it remains elevated in some branches of the supply chain such as ethnic food stores and restaurants (Di Muri et al., 2018), and in much of the world (FAO, 2018a).

Complex and fascinating, fisheries can be traced as far back as 42,000 years ago as evidenced by archaeological remains of pelagic species (O'Connor et al., 2011). Today, the ocean is considered one of the last heavily exploited wild frontiers. Sparked by

disastrous failures from the industry such as the collapse of the Grand Banks Atlantic cod fisheries, this trade has received growing attention from the media and is more often than not depicted as a destructive global enterprise putting marine ecosystems at risk, a message epitomized in Netflix's latest documentary on the matter, *Seaspiracy* directed by Ali Tabrizi (Tabrizi, 2021). Fishermen stand accused of using damaging harvesting methods, of driving stock depletion, and of proving careless with regards to by-catch. Successful books, such as *The End of The Line* by Charles Clover (Clover, 2013) later turned into a documentary (R. Murray, 2009), testify the interest that the general public has developed for issues surrounding fisheries. Today, experts diverge on their views of the state of world fisheries. Some prominent and mediatized fisheries biologists such as Daniel Pauly or Rashid Sumaila often paint a bleak picture of this industry and continuously denounce instances of ecological, economical, and social abuse in a trade that can sometimes be difficult to control and regulate (Pauly, 2019; Sumaila et al., 2020; Worm et al., 2006). Concepts coined in late 20th century by Daniel Pauly, such as the "shifting baseline" or "fishing down the food web" illustrate some philosophical and societal phenomenon described as having clouded the judgment of fisheries scientists when making management decisions. Other experts, such as Ray Hilborn paint a more optimistic picture and often point that, despite remaining issues in urgent need of solutions, many countries have reached sustainable levels of exploitation and numerous fisheries are on their way to recovery (Costello et al., 2020; Hilborn et al., 2020; Worm et al., 2009). Regardless of their stand, experts agree on one thing; poor management, corruption, fraud, and illegal practices still persist in this multi-billion dollar industry, and until these are controlled and managed accordingly, many stocks and species will remain at risk (Hilborn et al., 2020; Hilborn & Hilborn, 2019; Zeller & Pauly, 2019).

1.1.2 Illegal, Unreported and Unregulated fishing, traceability, and seafood mislabelling

Western countries are some of the most important consumers of seafood in the world

with North America ranking as the third highest per capita consumption in the world, and Europe the fourth highest (FAO, 2020). Both regions also happen to be considerable importers of seafood, with the European Union being the largest importer of seafood in terms of value in 2018, and the United States of America being the second largest importer in the world (FAO, 2020). The figures for these two regions alone highlight the importance of the international fish trade which in 2019, reached 8.55 million tonnes for the EU, second only to China, and 4.36 million tonnes for the United States (EUMOFA, 2020). Despite a strengthening of fisheries legislative frameworks in many countries, the increase in consumption, catch, aquaculture, and global trade are making traceability efforts incredibly challenging (Leal et al., 2015). Traceability, which can be defined as the ability to track back a product to its point-of-origin (Kehagia et al., 2007), is generally maintained through the use of labels and barcodes which should enable anyone throughout the supply chain to know where the product originated (e.g., what species it is, where it was caught, how it was processed, what methods were used to catch it, who caught it, etc...). Resulting from this complex trading system and operating both in domestic fishing grounds and in the high seas, fraudulent behavior in the fishing industry is an international phenomenon directly impacting transparency in the seafood industry and driving unregulated stock depletion (Sumaila et al., 2020). Some factors such as poor export and import regulations as well as flag of convenience, the practice of registering a vessel with a country that will often exhibit less severe regulations than that of the ship's owner, allow Illegal, Unreported, and Unregulated (IUU) fisheries to persist (Jacquet & Pauly, 2008). The practice of IUU fishing is first and foremost driven by economical incentives, leading a perpetrator to believe that fraud and its associated risks will result in economic gain (Gallic & Cox, 2006). Other fraudulent behaviours clouding traceability efforts, such as species substitution and deliberate mislabelling of seafood products can facilitate and conceal the phenomenon of IUU fishing.

Seafood mislabelling is a pernicious practice that greatly hampers our ability to sustainably manage resources by covering, hiding, laundering, and blatantly misleading consumers, managers, and enforcers. Though this phenomenon has recently received a

lot of attention, it was described as early as 1933 in the influential yet rather controversial book “100,000,000 Guinea Pigs: Dangers in Everyday Foods, Drugs, and Cosmetic” (Kallet & Schlink, 1933). Despite accusations of seafood mislabelling at the time, the issue received little attention and went rather unnoticed until much later in the century. The advent of genetics allowed to shed light on the sheer scale of this practice worldwide (FAO, 2018a) an it was reported in 2016 that over 50 countries worldwide were victim to this fraudulent practice (OCEANA, 2016) leading to important economic loss (Agnew et al., 2009).

The FAO labelling regulations are very limited, requiring that the country in which a product is processed be marked with the country of origin (*U.N. FAO. Codex General Standard for the Labelling of Prepackaged Foods, in CODEX STAN 1985*, 1985). Whereas some Western countries such as Canada have modest labelling requirements, others have devised their own, more stringent regulations (EU Regulations No 104/2000/EC, No 2065/2001/EC, and No 1169/2011). According to EU regulations for example, which are also some of the most rigorous, fresh seafood products must be labelled with the species scientific name, the catch location if caught in the Northeast Atlantic, the fishing gear, and the production method (caught versus farmed). In the USA a list of acceptable market names (FDA, 2020) serves as a labelling guide for the industry and certain fresh and frozen products must include Country Of Origin Labeling (COOL) (Country of Origin labeling for Fish and Shellfish, Code of Federal Regulations Title 7 Part 60). All of this information has the potential to be erroneous and mislabelled but for technical and technological limitations, the most prevalent type of mislabelling reported in the scientific literature pertains to species mislabelling (FAO, 2018a). Whereas some types of fraud such as the use of illegal fishing gear remain difficult to apprehend and report, others such as the catch location of a marketed specimens might become increasingly verifiable (Fields et al., 2020; Marko et al., 2011; Nielsen, Cariani, et al., 2012).

Some nomenclatures are particularly prone to mislabelling; in the USA, 70-80% of the fish sold as “red snapper” were revealed to be specimens from an array of different

species, many of which did not even belong to the Lutjanidae family (Cawthorn et al., 2018; Spencer et al., 2020); like “snapper”, the name “grouper” lands itself to being abused (Calosso et al., 2020). Though not always fraud, this kind of substitution and mislabelling is often motivated by a relatively high consumer demand for these types of fish and a low supply as many of these species will have been fished beyond sustainable levels. The issue of supply not meeting demand is a concern not only in terms of species substitution but also in terms of catch location substitution. If a stock has collapsed, fishers may either be tempted to keep fishing in that location and report the catch as coming from elsewhere or may catch fish elsewhere and declare it as coming from locations for which they still own quotas. Under such circumstances, the temptation to mislabel the catch location might come from various links in the supply chain including fish mongers, supermarkets, and restaurants. Though illegal fishing activities in regions with disputed geographical boundaries have been revealed using satellite imagery (Park et al., 2020), the extent of point-of-origin mislabelling of marketed products has rarely been documented due to a lack of simple and obvious tools to verify catch location.

1.1.3 Consequences of seafood mislabelling

Which ever form mislabelling and species substitution takes, it can result in severe negative consequences and in many cases, it is the result of fraud passable of fines and retribution (Martinsohn et al., 2019). The mislabelling of seafood products and the resulting money laundering has a direct economical impact (Agnew et al., 2009; A. Cohen, 1997; Donlan & Luque, 2019; Martinsohn et al., 2019). From an economical standpoint the worst scenario is when a low value species is being sold for a high value species (Calosso et al., 2020; Christiansen et al., 2018; Xiong et al., 2016). From an environmental standpoint, mislabelling may hide fraudulent practices from authorities and consumers opening the door to the trade of endangered or overexploited species, specimens from depleted stocks, and fish products that would otherwise be restricted to local consumption (Agnew et al., 2009; Miller & Sumaila, 2016). This in turn has the potential to fuel overfishing and unsustainable harvesting practices eventually leading stocks to collapse and species to extirpation (Fields et al., 2020).

Fisheries along with timber are some of the last heavily exploited wild resources and if traceability of these highly demanded products is compromised, so will the survival of the species that support this trade (’t Sas-Rolfes et al., 2019). In the grand scheme of things, reports of continuously mislabelled products have the potential to affect market seafood demand (Mariani et al., 2015). Mindful consumers with a desire to purchase sustainably harvested products are on the rise (Jacquet & Pauly, 2008) and may lose trust in the seafood industry should they realize that they were in fact deceived time and time again.

Importantly, some fish have the potential to hurt consumers if they are prepared improperly or consumed in high proportions. These fish are sometimes mislabelled and sold under another name, causing outbreaks of illnesses in a community. This was the case of the supposed Atlantic cod shipped to Hong Kong from Indonesia which proved to be a cargo of diarrhea inducing oilfish, *Ruvettus pretiosus* (Lam, 2007). In Chicago, two individuals were intoxicated in 2007 after consuming home-cooked puffer fish which was mislabelled as monkfish and contained tetrodotoxin (Cohen et al., 2009). More commonly, escolar (*Lepidocybium flavobrunneum*), which contains indigestible wax ester, gempylotoxin, is often mislabelled as white tuna and can be the cause of both intoxication and histamine poisoning (Feldman et al., 2005; Naaum et al., 2016).

Repeated occurrence of seafood poisoning has largely contributed to the development of better labels. Concerns have also been raised with respect to the mercury content of certain species such as yellowfin tuna or albacore tuna (Jacquet & Pauly, 2008). If these species are mislabelled and sold under the name of “skipjack tuna”, a species known to contain lower levels of mercury, this robs the consumer from the ability to make informed purchase decisions based on health considerations. Similarly, tilefish from the Gulf of Mexico is sometimes mislabelled which can be a concern given its high mercury content (Warner et al., 2012). Beyond issues of intoxication and mercury content, parasitic load of certain species can also cause food poisoning if those species are substituted and mislabelled (Williams et al., 2020).

1.1.4 Causes and incentives of seafood mislabelling

The mislabelling of a product indicates a problem somewhere down the supply chain, but it does not give any information as to where the problem stems from in this system or as to why this has occurred. The issue is complex and can be the result of a deliberate deceit for a variety of context dependent reasons like rampant poverty issues, but it can also be the result of lax fisheries management and weak regulations. Researchers have speculated upon the many reasons pushing parties throughout the supply chain to adhere to such fraudulent practices, but these are usually supported by anecdotal statements rather than quantifiable measures. Perhaps counterintuitively, a recent mislabel meta-analysis based on a few studies reporting product prices concluded that worldwide mislabelling is not driven by profit incentives (Donlan & Luque, 2019). The authors acknowledge however that there are many confounding variables that may have affected these "price-effect" results including fisheries regulation avoidance behavior, the observed stage of the supply chain (i.e. fishermen, fishmongers, and restaurant will have different incentives and motivations), and the substituted and substitute species, as well as their local management context. For example, in some conditions the cheaper Yellowfin tuna (*Thunnus albacares*) may be used to substitute the more valuable Atlantic Bluefin (*Thunnus thynnus*) tuna for profit, and in other conditions, the endangered Atlantic Bluefin tuna may be used to substitute the less threatened Yellowfin tuna in what is then called "reverse substitution" (Gordoa et al., 2017). Finally, high demand of some species in certain countries such as the Sablefish (*Anoplopoma fimbria*) in China, may lead to the trade and mislabelling of more pricey but less popular species such as the Patagonian (*Dissostichus eleginoides*) and Antarctic Toothfish (*Dissostichus mawsoni*) (Xiong et al., 2016).

In some cases, consumers disinterest with species labels, lack of seafood literacy knowledge, or poor local management procedures such as the absence of species label all-together can trigger fraudulent practices such as the trade of endangered species. Insidious mislabelling and smuggling of endangered species is also observed when

species are grouped for export, when they share similar physiological features, or when they are heavily processed, as is often the case in the shark product trade (Johri et al., 2019; Prasetyo et al., 2021; Sembiring et al., 2015). Again, though these are mere speculations, under such circumstances, it has been repeatedly suggested that this pernicious behavior is deliberate and driven by profit motives (Sembiring et al., 2015; Urbina, 2019). Whatever the motive is for fraudulent behavior, one cannot exclude unintentional substitution of species as a possible cause for mislabelling, driven either by carelessness, a lack of education and knowledge, or simply by an inability to care due to adverse circumstances such as poverty or modern slavery at sea (Urbina, 2019). All these factors affect price dynamics in unique and variable ways making it difficult to highlight and identify any one reason responsible for fraudulent behaviors in the seafood industry, and making evidence-based hypothesis difficult to formulate (Donlan & Luque, 2019).

1.2 DNA-based tools improving seafood traceability

1.2.1 Non molecular tools

Methods for evaluating seafood mislabelling are diverse and many have been used successfully to denounce fraud in the industry (FIGURE 1). DNA-based tools have largely been the preferred method for the identification of marketed fish, but stable isotope analysis have predominantly been used to determine seafood geographical provenance (Gopi et al., 2019). DNA is particularly suited for the identification of species due its mode of inheritance and to the existence of comprehensive databases indexing hundreds of thousands of species (Rasmussen Hellberg & Morrissey, 2011). For processed samples, it has been preferred over taxonomic methods which require certain elements of the specimens to be intact (fins, otoliths, skin, etc...). Other methods such as physicochemical analysis, sensory analysis, and rheological offer means to identify species but require the use of highly specialized and expensive instruments. Hyperspectral imaging techniques offer a relatively quick and cheaper option (Qin et al., 2020) but would require the development of a species database and the purchase of a

hyperspectral instrumental system and are therefore infrequently used in seafood species identification. With regards to provenance testing a variety of bio-tracers have been explored such as fatty acid profiling and elemental profiling, but stable isotope analysis has largely been the preferred method (Gopi et al., 2019). Stable isotope analyses are particularly useful as they reflect the foraging geographical location of an individual and offer an insight on trophic levels. Their advantage and disadvantage for the provenance testing is compared over molecular tools and discussed extensively in Chapter 4.

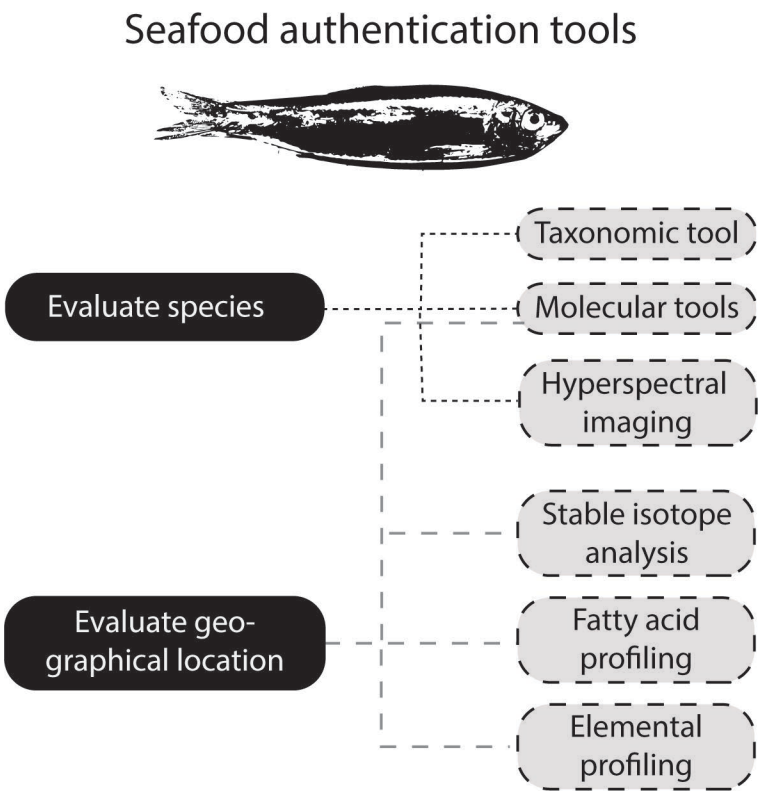


FIGURE 1.1 Authentication tools for identification of seafood species and geographical catch location.

1.2.2 The advent of the barcode of life

Though taxonomic identification of species is invaluable, it can sometimes prove inefficient, laborious, and greatly expertise dependent. It has also occasionally led to erroneous conclusions and some genetic studies are now shattering decades worth of taxonomic-based studies by redefining entire species geographical distribution (Choquet

et al., 2017). With the advent of molecular technologies, Hebert et al. (Hebert, Ratnasingham, et al., 2003) set out on the ambitious mission to develop a system that would identify any animal species on the planet based on a single short fragment of the genome. Hebert et al. (2003) argued that the mitochondrial DNA (mtDNA) is best for species identification as it does not contain introns (non-coding segments), it seldom goes through recombination, and it is also inherited via a haploid mode (just from the mother). Additionally, mitochondria being present in high numbers in a single cell, there are substantially more copies of mitochondrial DNA than nuclear DNA facilitating laboratory procedures (Antoniou & Magoulas, 2014). Any gene in the mtDNA could technically have worked for microgenomic species identification, but cytochrome c oxidase subunit 1 (COI) advantageously has a very robust universal primer region for animals.

Today, the mtDNA COI region is referred to as the barcode region for species identification and has been described for more than 200 000 species. These barcodes are available for Basic Local Alignment Search Tool (BLAST) analysis on the barcode of life data system (http://www.boldsystems.org/). Following the successful implementation of the COI barcode species identification system by Hebert et al. (2003), the international community took the initiative to catalog all fish COI profiles into a library: The Fish Barcode of Life (FISH-BOL; www.fishbol.org) on-line resource (Ward et al., 2009). Once completed, such resource should allow for the identification of any fish species no matter its life-stage, degradation level, or how cryptic it might be. Taxonomic identification being particularly difficult for processed marketed fish, the use of mtDNA has successfully led studies to highlight the rate of mislabelling around the world (Hanner et al., 2011; Miller & Mariani, 2010; Ogden, 2008; Yancy et al., 2008).

Though mtDNA can be used as a practical taxonomic tool, it also fails to be effective in some cases such as with the distinction of hybrids, a problem linked to the fact that mtDNA is maternally inherited. mtDNA can also sometimes fail to identify species or on the contrary can indicate phylogenetic variations when nuclear DNA does not suggest any.

These imperfections have led bodies of researchers to suggest that despite its apparent advantages mtDNA should not be viewed as the panacea of phylogenetic and taxonomic inference (Rubinoff & Holland, 2005). In the context of seafood species authentication however, it has largely proven to be a useful and reliable tool.

1.2.3 Overview of DNA-based tools in seafood traceability

Advances in genetic technologies have permitted DNA-based methods to inform fisheries conservation management (Bernatchez et al., 2017) and have recently been successful in denouncing the ubiquitous issue of seafood mislabelling (Wong & Hanner, 2008; Rasmussen & Morrissey, 2009; Miller & Mariani, 2010). Increasing scientific interest and publications examining this problem prompted extensive media coverage and consumer awareness initiatives (Mariani et al., 2014). In Europe, studies highlighting the extent of mislabelling in the market are likely at the root of governmental actions and the implementation of more stringent regulations. This attempt to tackle the issue has led to a surprising decrease in mislabelled seafood in Western Europe (Mariani et al., 2015). As discussed above however, this amelioration is not uniform across the European market (Di Muri et al., 2018; Vandamme et al., 2016) or across commercial species (Cawthorn et al., 2018), and it remains quite unique in the face of a world now exhibiting the highest recorded seafood demand (Fields et al., 2020). In Europe and North America, independent organizations such as the Marine Stewardship Council (MSC) offer voluntary sustainability standard audits via which fisheries can ask to get certified and prove their reliability to consumers. Others, such as the Marine Conservation Society have developed a rating scheme via which they can recommend or discourage the consumption of certain fish species. Despite these incentives, IUU fishing contributes to a fifth of total seafood harvest and continues to represent a global challenge (Agnew et al., 2009; FAO, 2020; Sumaila et al., 2020). In order to maintain their integrity, eco-labels must be able to evaluate whether there is any risk of labelling fraud stemming from the product they certify. DNA-based tools could offer a solution in verifying the authenticity of certified seafood products both with respect to species substitution and the geographical point-of-origin.

1.2.4 Quick DNA-based species identification tools

DNA-based methods have proved to be an invaluable tool to expose seafood mislabelling worldwide and have been proposed as a regulatory and prosecution tool. There are few and far between examples of the use of DNA-based methods in an enforcement context, often leading to prosecution and fining of the perpetrators (Martinsohn et al., 2019). Perhaps one of the major drawbacks of genetic methods is their perceived complexity and the relatively slow result rate from analyzed samples. Sanger sequencing, which is commonly used for the authentication of seafood product, often requires the externalization of the analysis by a third party laboratory to process post-PCR product (Rasmussen Hellberg & Morrissey, 2011). This can be time consuming particularly in countries where sequencing devices are not readily available, and the samples need to be shipped internationally. Sanger sequencing in the context of seafood fraud has been useful as an exploratory research tool but is maladapted as a systematic screening tool. In the context of enforcement, governmental organizations, stakeholder, and NGOs are repeatedly asking for tools that are quick and easy to use, requiring little to no expertise and allowing for the processing of specimens on site. Quick and cheap identification methods do exist, though there is currently no standardized protocol used across all aquatic species. Restriction fragment length polymorphism (RFLP) is a method that has been commonly used for the identification of species based on predetermined short target DNA fragments which are then visualised using gel electrophoresis (Handy et al., 2017; Silva & Hellberg, 2021). Though some rapid RFLP assays have been developed commercially by Agilent Technologies (Santa Clara, CA), they require the use of an Agilent Bioanalyzer Instrument which restricts the consumer to using lines of products developed by the company and which limits them to the number the species for which the assay was designed (Handy et al., 2017). In addition to these limitations, the method requires use of specific restriction enzymes for different species depending on the RFLP polymorphism, generates ambiguous pattern for related species, and ultimately is too cumbersome for field work. There are however, some successful enforcement stories that have been made possible thanks to robust and quick portable DNA-based identification tools, such as

species specific quantitative-PCR (qPCR) assays (Cardeñosa et al., 2019). High resolution melt curve (HRM) which relies on the denaturation of DNA during an increase in temperature in the presence of highly concentrated saturation dyes can provide unique melting curves for the authentication of species. It has been successfully used in the detection of food fraud (Druml & Cichna-Markl, 2014; Silva & Hellberg, 2021). HRM has been developed for the detection of various gadoids (Shi et al., 2020), shrimps (Fernandes et al., 2017), and mussels (Quintrel et al., 2021) and proves robust against variations in single nucleotides.

Other emerging third generation sequencing technologies such as nanopore sequencing may prove useful in the quest to authenticate specimens quickly and on site (Menegon et al., 2017; Pomerantz et al., 2018), but may be cumbersome in terms of the required bioinformatic expertise. Recently, closed-tube barcoding was highlighted as a potential quick and universal method for fish species authentication (Naaum et al., 2021 - See Appendix). Though promising, the technology is relatively new and needs more testing to prove reliable and easy to use in the context of seafood verification.

1.2.5 DNA-based tools for population assignment

1.2.5.1 DNA-based tools for population assignment

Due to lasting reproductive isolation between taxa, DNA sequences diverge between lineages and can easily be used for species discrimination and identification (Hebert, Cywinska, et al., 2003). On the other hand, populations from a given species need to be sufficiently reproductively isolated for genetic variants to be detectable (J. R. Freeland & Petersen, 2011; Waples & Gaggiotti, 2006). Population boundaries can be difficult to define, and their degree of isolation will vary from panmixia to complete isolation (Waples & Gaggiotti, 2006).

The choice of molecular marker is incredibly important when estimating genetic diversity and is a crucial aspect of population genetic methods. Mutation rates will differ substantially between markers making them more or less adapted to population structure analysis depending on the question of interest. Mitochondrial DNA cytochrome b and

cytochrome oxidase 1 (COI) have been used extensively for species identification due to their fairly “low” mutation rate making them poor marker choices for population structure studies. Markers can also be chosen based on their functionality, i.e. whether evolutionary constraints affect them or not (Mariani & Bekkevold, 2014).

Nuclear microsatellites and single nuclear polymorphism (SNP) markers have been used quite extensively for population genetic studies. Though microsatellites are non-coding markers, leading them to mutate at a faster rate than coding alleles, they may still be under the pressure of selection when occurring near coding regions. SNPs are variations in a single base pair that are scattered across the genome and that can be part of both coding and non-coding regions. SNPs seem to allow for a greater resolution than other markers as thousands of them can be screened simultaneously, and ‘outlier’ loci under diversifying selection can be high-graded to maximise genetic divergence between populations (Mariani & Bekkevold, 2014; Nielsen, Cariani, et al., 2012).

Due to their commercial importance and heavy exploitation, salmonids have been particularly well studied with methods ranging from tagging to genetics as means to investigate population structure. Various extensively used genetic markers such as microsatellites and more recently SNPs have revealed that salmonids are distinctly and hierarchically structured and that some species like Atlantic salmon exhibit patterns of isolation by distance (Glover et al., 2012). The use of genetic markers have proved successful in answering a great diversity of population-related questions in salmonids such as evaluating the level of introgression between wild salmon and farmed escapees (Glover et al., 2012; Harvey et al., 2015), defining management units (Palsbøll et al., 2007) stock boundaries (Gilbey et al., 2018), and reproductive isolation (Ramstad et al., 2004; Wennevik et al., 2019), but also understanding the impact of climate change on salmon population structuring (Horreo et al., 2011). These studies have not only proved that genetics can be successfully used to investigate questions pertaining to population genetics in fish but have also helped develop the field further.

1.2.5.2 Traits affecting population genetics in the marine environment

In marine ecosystems, lack of absolute physical barriers, complexity of oceanographic variables and life history traits, and high demographic connectivity have long challenged the study of genetic divergence among populations. Currents, temperature gradients, depth, and bathymetry form a multivariate framework of environmental factors that dictates connectivity between populations. Adding to these complex abiotic elements, species life history such as migration patterns, egg and larvae buoyancy, larval duration, and diel vertical migration can lead to dispersal and increase connectivity between populations (Bohonak, 1999; Stepien, 1999; Young et al., 2015). Though this cannot be applied to all marine species, certain life history traits such as high fecundity, high larval dispersal, and high migration patterns, have facilitated near-panmixia among populations of certain species (Palumbi, 2003).

The concept of seascape genetics considers that both biotic and abiotic parameters can affect population genetic structure. Seascape genetics is a interdisciplinary approach to exploring and explaining population structure and takes advantage of our increased understanding of physical oceanic processes and bathymetric profiles, as well as life-history traits (Selkoe et al., 2008). Such resolution is important to understand the drivers of population structure but will also point to the fact that no single driver affect population structure but rather a combination of them. Whereas biological traits usually pertain to species (e.g. migration patterns, dietary habits, larval duration, etc...), abiotic factors are locally measured parameters that can differ from population to population.

1.2.5.3 Molecular tools for provenance testing: success and failures

If the point-of-origin of a marketed product is to be determined in the context of regulatory frameworks, the question of interest will not so much be “where does this product come from” but rather “does this product come from the area reported on the label?”. When it comes to DNA-based technologies, identifying population structure is a necessary first step towards developing provenance testing tools. If a given species does not exhibit population genetic structure, molecular tools will be ineffective. If, however, certain

markers prove successful in discriminating between populations of interest, an assignment probability can then be generated. This allows to evaluate the accuracy with which a specimen can be traced back to its population of origin. High assignment percentages provide high point-of-origin certainty and create a realm of opportunity for molecular technologies as tools for regulatory purposes.

Despite this promising framework and although assignment percentages have now been assessed for a number of commercially important species (FishPopTrace, 2013), management and regulatory entities have failed to demonstrate interest in using these tools, aside for some very isolated, and successful, cases (Martinsohn et al., 2019; Nielsen, Hemmer-Hansen, et al., 2012). Though some have proposed that the expenses related to genetic analysis are a potential drawback, Martinsohn et al. (2019) unambiguously demonstrated that the financial benefit due to monetary penalties largely outweigh the costs of a DNA-based forensic analysis for origin assignment. Perhaps the reluctance to use molecular tools comes from the perceived complexities and possible uncertainties associated with their use rather than their cost. Indeed, a major obstacle to genetic tools for point-of-origin assignment investigations is the current lack of consistency in terms of marker choice and methodology. Inconsistencies in marker choice and methods may also complicate the development of complete and standardised reference baselines for assignment. Another major challenge to overcome with the use of genetic methods for point-of-origin studies simply pertains to the discrepancies that still exist between geographical management zones and biological population boundaries (Reiss et al., 2009).

1.3 Aims and objectives

Seafood traceability has greatly improved over time, but fraudulent practices are still rampant in the fisheries industry leading to malpractice such as mislabelling and confused or lacking product transparency. This thesis examines some of the societal factors that may affect seafood traceability and explores DNA-based methods that have the potential to greatly improve the continuous and regular monitoring of transparency and traceability along the supply chain. Each chapter is dedicated to a given driver which might allow mislabelling to persevere (i.e. lack of consumer knowledge, shortcomings of species identification methods, absence of a framework for the use of point-of-origin detection tools) and explores some of the associated solutions that could help strengthen the monitoring of seafood products, verify compliance, and tackle fraud in the seafood industry. Through this journey, the thesis explores ways to improve seafood traceability by addressing the following objectives:

- To understand and evaluate consumers' level of seafood literacy (**Chapter 2**)
- To explore the use of a fast and portable DNA-based technology for the on-site identification of fish species (**Chapter 3**)
- To assess the general effectiveness of using genetic tools for the identification of fish geographical provenance and design a framework for geographical assignment applicable to any commercial species and marketed specimens (**Chapter 4**)
- To practically test a DNA-based provenance testing method in the context of an international market study of Atlantic cod (**Chapter 5**)

Chapter 2

Fish out of water: Consumers' unfamiliarity with the appearance of commercial fish species.

This chapter was published in *Sustainability Science*

Authors: **Marine Cusa**, Luara Falcão, Joseph De Jesus, Cristina Biolatti, Lancelot Blondeel, Fiona Bracken, Lisa Devriese, Sandra Garcés-Pastor, Stylliani Minoudi, Chrysa Gubili, Pier Luigi Acutis, Stefano Mariani

Note: The data collection was performed prior to Marine's engagement in this study. Marine Cusa was responsible for curating and analyzing the data, designing all figures and diagrams, and writing the manuscript

2.1 Abstract

Seafood labels play an increasingly key role in assisting consumers in purchasing processed and featureless fish products. While informed purchasing choices are typically influenced by traceability and labelling awareness, they also depend on the consumer's ability to identify and discriminate the fish species available on the market, which to date remains notably unexplored. We asked 720 people across six European countries to identify pictures of six fish species commonly sold in Europe. We reveal that European citizens have a poor understanding of the appearance of the fish they consume (overall 30.19% correct identification), with British consumers performing the poorest and Spanish ones doing best. We noted cultural association with some species, whereby the most regionally consumed fish are more easily recognized. We argue that despite recent improvements in technological solutions, stakeholder dialogue, and policy implementation, seafood market transparency will remain open to malpractice until consumers restore connection with their food.

Keywords

Sustainable fishing · Seafood traceability · Mislabeling · Environmental awareness · Food literacy · Marine Conservation

2.2 Introduction

Seafood product substitution and lack of transparency in the supply chain opens the door to the trade of endangered species, to unsustainable aquaculture and fishing practices and to the depletion of stocks, while exposing consumers to health and safety risks and enabling fraudulent business (A. Cohen, 1997; Fields et al., 2020; Jacquet & Pauly, 2008). While neglect or genuine mistakes are sometimes the root cause of seafood mislabeling, deliberate fraudulent behavior driven by the appeal of economic gain or by the trade of products derived from illegal, unreported, and unregulated (IUU) fishing will also result in instances of mislabeling (Calosso et al., 2020; Cawthorn & Mariani, 2017; Donlan & Luque, 2019). The advent of DNA-based tools has shed light on the worldwide scale and pervasiveness of seafood mislabeling, in some cases escalating tangible change through improved governance and media attention (Mariani et al., 2014). This is particularly true for Europe, where the mislabeling of certain species has been shown to substantially decrease as a response to public awareness and improved legislation (Mariani et al., 2015). Yet, it remains unclear whether these rapid improvements can be sustained over the longer term. Consumers knowledge on fisheries resources has been credited with a role in fostering seafood sustainability (Olson et al., 2014), but virtually nothing is known about consumers' familiarity with the fish they eat, a concept that we hereby term 'seafood literacy'.

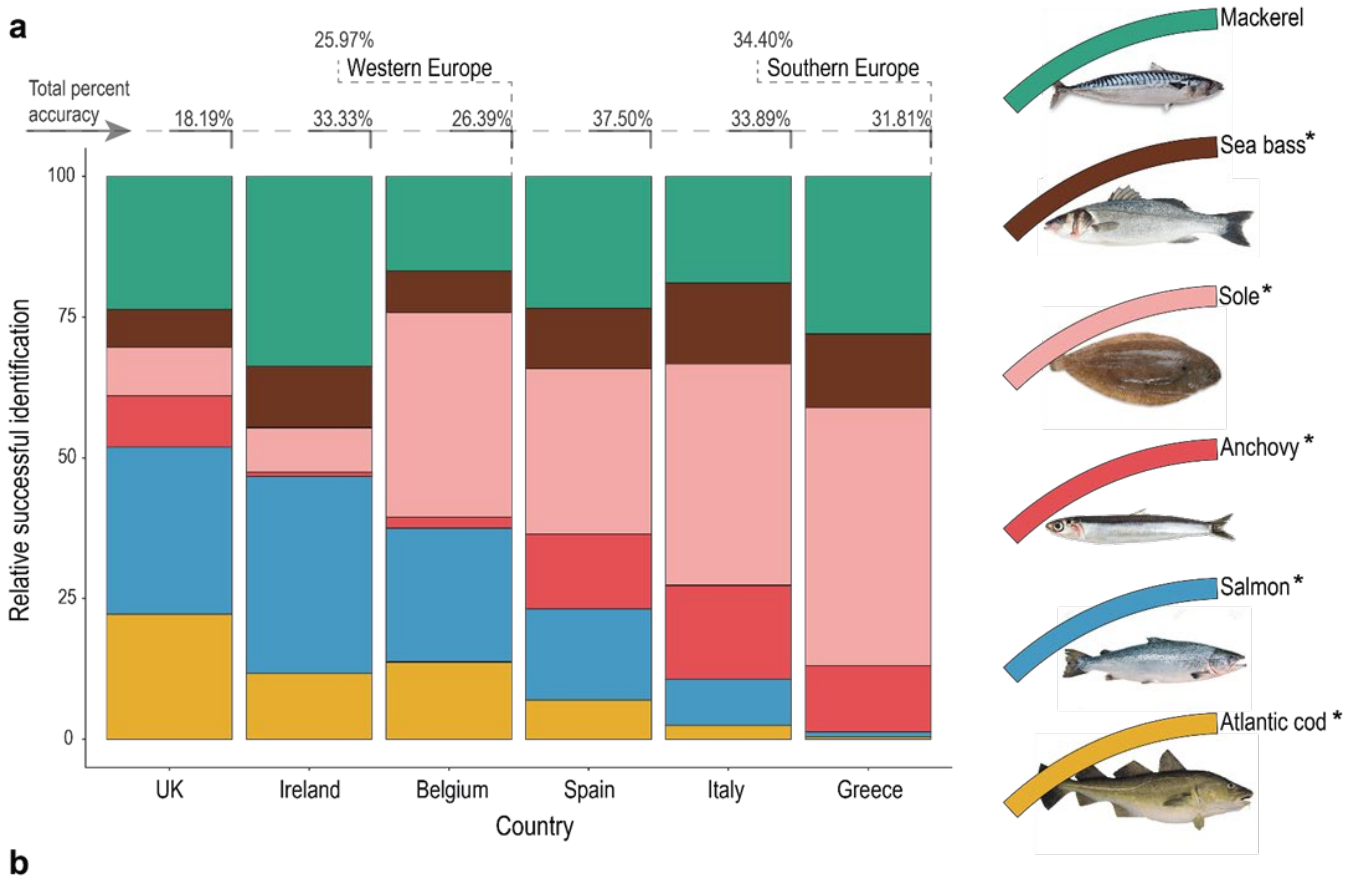
Though there is no set definition for the concept of Food Literacy, the majority of experts view it as the knowledge required for consumers to make informed purchasing and feeding choices with regards to personal health, environmental impact, and ethical standards (Bellotti, 2010; Perry et al., 2017; Vidgen & Gallegos, 2014). Species literacy, a recent concept coined by Hooykaas et al. (2019), encompasses one's knowledge on specific species including the ability to recognize species visually. In fact, the sole ability to identify and name species has previously been associated with greater levels of affinity, respect, and appreciation (Mohneke et al., 2016; Schlegel & Rupf, 2010). Given the great diversity of species on the seafood market and their diverging life histories, ecological

roles, and conservation status, consumer ability to recognize species is a particularly relevant and empowering aspect of seafood literacy (Gaviglio et al., 2014). Fish products are too often considered as a homogenous commodity, grouped under the term 'Fish' (Gaviglio et al., 2014), or frequently aggregated in generic categories or 'umbrella terms' (Cawthorn et al., 2018; Griffiths et al., 2013). The use of hypernyms in the seafood industry or deliberate mislabeling practices can hide the trade of vulnerable species, or of species prone to IUU fishing (Calosso et al., 2020; Cawthorn & Mariani, 2017) and can lead to the oblivious consumption of fish from poorly managed stocks (Kroetz et al., 2020). A limited ability to recognize and distinguish between species can potentially result in a general lack of concern for marine biodiversity (Balmford et al., 2002; Schlegel & Rupf, 2010) thereby encouraging fraudulent behaviour and allowing damaging practices, such as fish substitution, to persist.

Globalisation and technological advances, such as improvements in freezing abilities, have also changed our relationship with food and have led to increased commoditization of seafood (Anderson et al., 2018). A 350% rise in seafood demand since the mid-1970s (FAO, 2017) and the resulting increase in seafood commoditization is responsible for a decreasing demand in species-specific products and for a growing tolerance in the substitution of species within key groups (Anderson et al., 2018) such as white fish, tuna, salmon, etc... These factors bear a tremendous impact on consumption habits and have emotionally and physically detached urban consumers from the source of the food they purchase (Bellotti, 2010; Vileisis, 2008).

Given the importance of consumer knowledge in an increasingly complex seafood market, studies have sought to identify how consumers react to labelling and traceability tools (Altintzoglou & Nøstvold, 2014; Rodriguez-Salvador & Dopico, 2020; Vitale et al., 2020), yet to our knowledge, none has quantitatively assessed buyers' familiarity with the appearance of the fish they purchase. In this study, we first assessed consumers' response accuracy with respect to the identification of commonly consumed fish species and explored how this might differ between countries and/or regions. We then explored

regional patterns of inaccuracy with a focus on the richness and diversity of wrong answers. We conducted this evaluation of consumers' ability to identify widely available seafood species in urban centers of six European countries. This allowed us to encompass comprehensive cultural range and urban seafood consumption habits, in 'Southern' and 'Western' European countries, as defined by the EU Commission map (EUMOFA, 2017), with the former typically having higher seafood consumption rate and a closer relationship with fishmongers, and the latter consuming less seafood and showing greater reliance on processed seafood available in supermarkets (EUMOFA, 2017, 2019).


a
25.97%
Western Europe
34.40%
Southern Europe
Total percent accuracy
18.19%
33.33%
26.39%
37.50%
33.89%
31.81%
Relative successful identification
100
75
50
25
0
UK
Ireland
Belgium
Spain
Italy
Greece
Country
Mackerel
Sea bass*
Sole*
Anchovy*
Salmon*
Atlantic cod*

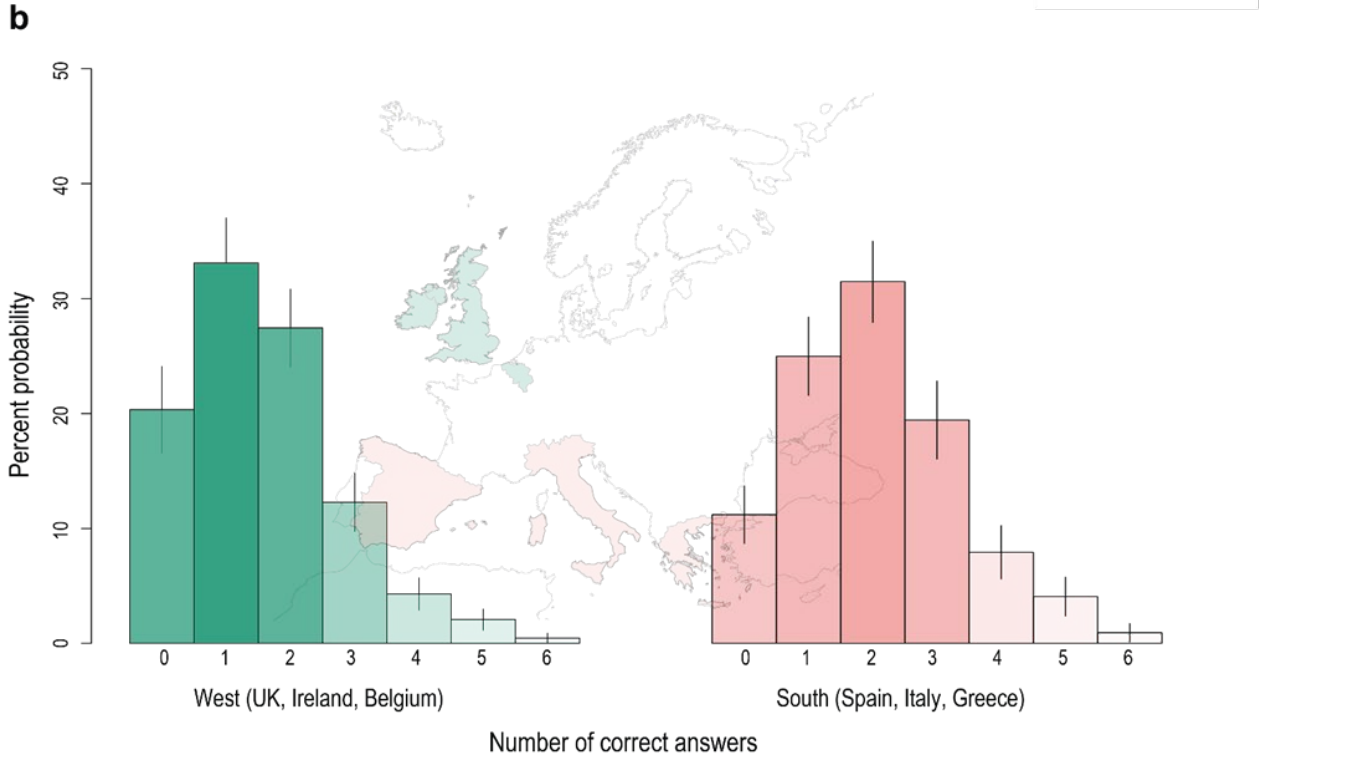

b
Percent probability
50
40
30
20
10
0
0 1 2 3 4 5 6
West (UK, Ireland, Belgium)
0 1 2 3 4 5 6
South (Spain, Italy, Greece)
Number of correct answers

FIGURE 2.1 (a) Relative successful identification of each of the six fish species for each of the six countries. The stacked bar graph only depicts correct identifications and the total percent accuracy is shown above the stacked bars for each country and for both regions. The bar graph segments represent how much percentage can be attributed to each species out of the total number of correctly identified fish. For instance, out of 31.81% identification success in Greece, about half is due to sole. Consumers' correct answers were compared between regions for each species independently, and the star * indicates that identification accuracy for that species was significantly different between Western and Southern European regions as tested via categorical Chi-square tests. The photos presented to the participants are depicted in the legend along with the species common name. **(b)** Consumers probability of falling into categories 0 to 6 (0 = zero fish correctly identified and 6 = six fish correctly identified) for northern countries (United-Kingdom, Ireland, and Belgium) and southern countries (Spain, Italy, and Greece) as predicted from a cumulative link model (clm) fitted values. The error bars correspond to the ±1.96 standard error for the mean prediction. The ordinal regression analysis indicates that regions differ significantly from each other in terms of the number of correct answers ($z = -5.21$, $p = 1.89 \times 10\text{-}7$ See Table S1 for threshold coefficients).

2.3 Methods

Data collection. We recorded consumers' ability to visually identify common fish species across six countries: Belgium, Ireland, United Kingdom ('Western Europe'), Greece, Italy, Spain ('Southern Europe'). A total of 720 consumers (120 per country) were interviewed in general retail areas of the city centers of Ostend, Dublin, Manchester, Thessaloniki, Turin, and Barcelona, between October and December 2016. City centers were chosen for this study to avoid socio-economic biases linked with certain neighborhoods. Consumers were chosen randomly in the crowd and were not purchasing seafood at the time of the interview. Consumers were asked about their professional background, age, and seafood consumption habits. Participants were also asked to identify a set of six fish species which were presented in the form of pictures. As illustrated in **FIGURE 2.1.a**, the photos illustrated the side view of a whole specimen and the species included four major fisheries resources of European waters: Atlantic cod (*Gadus morhua*), Atlantic mackerel (*Scomber scombrus*), European anchovy (*Engraulis encrasicolus*), and Common sole (*Solea solea*), and the two top mariculture species produced in Europe: Atlantic salmon

(*Salmo salar*) and European seabass (*Dicentrarchus labrax*). Correct, incorrect and void answers were recorded and standardised across countries using official English language denominations. Vernacular versions of fish names and their regional variants across countries (i.e. 'mackerel' for Atlantic mackerel, or the Italian 'branzino' for European seabass) were all accepted as correct responses.

Statistical analysis. All statistical analyses were performed in R (R Core Team, 2019; https://www.r-project.org/). The regional identification accuracy for each individual fish species (i.e. Atlantic cod, Atlantic mackerel, European anchovy, Common sole, Atlantic salmon, and European seabass) presented to consumers was tested using a Chi-square tests. For each species, we expected that identification accuracy differs significantly between WEU and SEU regions. A Cumulative Link Model (clm, also known as Ordinal Logistic Regression) was used to test whether the number of correct answers from participants (0-6) differed significantly between regions. clm is a powerful non-parametric test well suited for data with ordinal dependent variables such as these. We used the package "ordinal" to generate the model (Christensen, 2019). To test whether percentage accuracy of responses differed significantly between countries, we used a non-parametric Kruskall-Wallis test followed by a Dunn post hoc test as our data violated parametric test assumptions. To do so, we used the package "FSA" (Ogle et al., 2020). Given similarity of response between countries of a given region and the observed differences in accuracy between regions, we here extrapolate the results from each city to the country they belong to. For regions, parametric test assumptions were met, permitting us to evaluate the effect of region on percent accuracy of response using a student's *t*-test.

To evaluate how wrong answers for each of the six species differed between regions and countries, we used two diversity indices: species richness *R*, and the Shannon diversity index, *H'*. The richness of wrong answers allowed us to quantify the total number of wrong species participants suggested for each country and region, whereas the Shannon diversity index allowed us to account not only for the number of wrong species but also the frequency of each wrong answer. Species richness was calculated as the

total number of species wrongly guessed and the Shannon diversity index was calculated using the following formula:

$$H' = -\sum_{i=1}^{R} p_i \ln p_i$$

where $p_i$is the proportion of wrong answers belonging to the *i*th species. A one-way ANOVA was used to determine whether the richness and Shannon diversity means differed significantly between regions and countries, and a post-hoc Tukey HSD test was performed for pairwise comparison among tested countries. For most graphical representations, the R package "ggpubr" was used (Kassambara, 2020) and figures were processed and assembled in Adobe Illustrator (Adobe Inc., 2019). The chord diagrams of the diversity of wrong answers were generated using the Circos software (Krzywinski et al., 2009).

2.4 Results & Discussion

2.4.1 Seafood consumption and unfamiliarity with fish appearance

Consumers unfamiliarity with the appearance of common fish species was flagrant, with an average of 30.19% identification accuracy across all countries. A Kruskal-Wallis Test indicated that percentage accuracy of response was significantly different (Chi square = 61.47, $p < 0.001$, df = 5) among the six countries (see **TABLE 2.1** for KW post-hoc Dunn's test), and a *t*-test indicated that region provenance of consumers affects the percent accuracy of their response [$t(718)=5.26$, $p < 0.001$]. Western European (WEU) countries did relatively worse, with the United Kingdom recording the lowest identification accuracy of 18.19%, followed by Belgium with 26.39%, while Southern European (SEU) countries performed slightly better on average, with values ranging between 31.81% (Greece) and 37.50% (Spain) (**FIGURE 2.1.a**). Significant regional differences in species-specific identification success were also apparent for five out of the six species we presented to

consumers. Unsurprisingly, cod and salmon, traditional mainstays of cold temperate waters, accounted for most of the successful identifications in the UK and Ireland (where they are also most consumed, EUMOFA, 2017), while warmer-water species such as anchovy, seabass and sole accounted for most of the successful identifications in Italy and Greece (**FIGURE 2.1.a**), suggesting that identification ability is intricately linked to traditional culinary habits and to the historical availability of species in the countries' waters and markets.

TABLE 2.1 Post-hoc Dunn-test displaying Z test statistic and adjusted p-values for pairwise comparisons of percentage accuracy of correct answers from consumers among six European countries.

Comparisons	Z	Adjusted p-value	
Belgium - Greece	**-2.4373**	**0.0074**	*
Belgium - Ireland	**-2.9805**	**0.0014**	*
Greece - Ireland	-0.5432	0.2935	
Belgium - Italy	**-3.169**	**0.0008**	*
Greece - Italy	-0.7317	0.2322	
Ireland - Italy	-0.1885	0.4252	
Belgium - Spain	**-3.6658**	**0.0001**	*
Greece - Spain	-1.2285	0.1096	
Ireland - Spain	-0.6853	0.2466	
Italy - Spain	-0.4968	0.3097	
Belgium - UK	2.7403	0.0031	
Greece - UK	**5.1776**	**0.0000**	*
Ireland - UK	**5.7208**	**0.0000**	*
Italy - UK	**5.9093**	**0.0000**	*
Spain - UK	**6.4061**	**0.0000**	*

When countries were separated into Western and Southern clusters, based on the EU Commission map (EUMOFA, 2017), consumers from Southern EU countries revealed finer identification skills than consumers from Western EU countries. The ordinal regression analysis indicated that regions differed significantly from each other in terms of the number of correct answers ($z = -5.21$, $p < 0.001$ see **TABLE 2.2** for threshold coefficients) and overall, greater identification accuracy can also be seen in the distributions of

probabilities in (**FIGURE 2.1.b**). Although regional discrepancies had previously been highlighted in terms of mislabeling rate and consumer awareness (Bréchon et al., 2016; Miller, Jessel, et al., 2012), which likely arise from the interplay of several cultural, demographic, and regulatory differences, the extent and implications of seafood literacy in the population had up to this point been largely neglected. Seafood consumption habits differ greatly among the countries investigated (K. Brunsø, 2003; EUMOFA, 2017, 2019), which also may affect fish species recognition performance. Spanish respondents demonstrated the highest level of identification success, and over 85% indicated that they consumed seafood at least once a month, a rate higher than in any other surveyed country (**FIGURE 5.2.a**) and consistent with previously reported data (EUMOFA, 2017). The UK and Greece on the other hand currently have the lowest per capita consumption of seafood among the six countries, which was reflected in the responses to our survey (**FIGURE 5.2.a**); yet, identification accuracy in Greece is significantly higher than in the UK (Kruskal-Wallis post-hoc Dunn's test $p < 0.001$, **TABLE 2.1**), likely due to the intersection of another factor: the way fish is traditionally presented to the consumer at retail level. SEU citizens are typically exposed to a greater diversity of species, displayed whole in most food stalls (**FIGURE 5.2.b**), which has arguably led to greater familiarity with the appearance of fishes in southern countries, while WEU fish markets generally offer a relatively narrow range of species, often already filleted (**FIGURE 5.2.c**).

TABLE 2.2 Cumulative Link Model threshold coefficients as calculated by the clm package in R for ordered data from participants (total number of species' photographs correctly identified per participant: 1,2,3,4,5, or 6) and compared between regions [test estimate = -0.7055, standard error: 0.1354, z-value = -5.21, Pr(>|z|) = 0.000000189].

Threshold coefficients	Estimate	Standard Error	z value
0|1	-2.0706	0.1292	-16.03
1|2	-0.5674	0.1015	7.102
2|3	0.7376	0.1039	7.102
3|4	1.9084	0.1369	13.94
4|5	2.9458	0.2036	14.47
5|6	4.6687	0.4519	10.331

The trends illustrated above likely represent cultural and market-based legacies of a recent past where seafood supply chains were less globalised (see D. D. Miller et al., 2012 for a case study). More recently, the appetite for fast and easy to prepare meals (EUMOFA, 2019), especially among urban-dwelling younger generations (Brunsø et al., 2009), is partly responsible for motivating the retail sector to prioritize processed products over fresh ones, favoring supermarkets to the detriment of fishmongers. Between 1988 and 1995 the largest retail chains in the United Kingdom, for instance, have gone from selling 31.5% to 60.9% of all the fresh fish (A. D. Murray & Fofana, 2002). Fishmongers and supermarkets adopt different product presentation strategies, and whereas fish are more likely to be laid fresh and whole on a stall in fishmonger stores, they tend to be presented as processed products in shelves or freezers in supermarkets (A. D. Murray & Fofana, 2002). Packaged and labelled seafood products often contain processed fish that lack morphological characteristics, forcing the consumer to rely on labels for product information, including species identification (De Almeida et al., 1997). The explosive increase of mariculture, which jumped from 5 million tons in 1988 to 31 million tons in 2018 (FAO, 2020) also likely contributed to the gradual dilution of regional associations between people and fish in Europe. This is epitomised by Atlantic salmon which jumped from a global production of 38 thousand tons in 1985 to 2 million tons in 2016 (Jones, 2004) and which was the second most consumed fish after tuna and the most consumed farmed species in the EU in 2018 (EUMOFA, 2020).

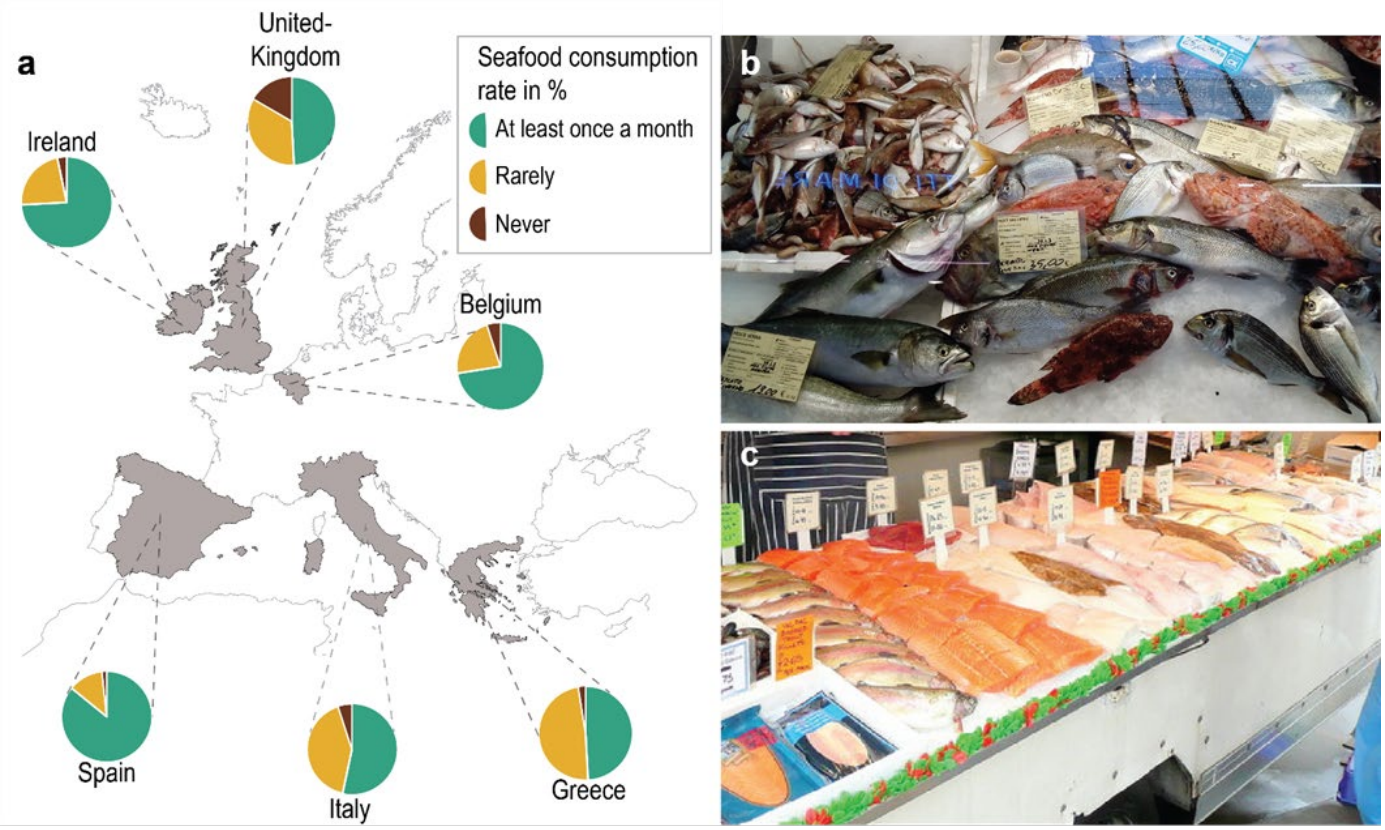


FIGURE 2.2 (a) Seafood consumption rate in percentage for a total of 720 participants from six European countries. **(b)** an example of a Mediterranean seafood stall (Livorno, Italy); **(c)** a northern example (Hereford, UK).

2.4.2 Consumers' perception of seafood

When asked to identify fish pictures, participants from WEU and SEU countries not only differed in terms of the accuracy of their response but also in terms of the richness and diversity of answers they provided. We harnessed the information contained in the incorrect answers to glean insights into the consumers' perception of the nature and diversity of seafood. An ANOVA indicated that means were significantly different between countries for both species richness [$F(5,30)=11.8$, $p < 0.001$] and Shannon-Weiner index [$F(5,30)=4.08$, $p < 0.001$]. Mean species richness and mean Shannon-Weiner diversity was also calculated for each region, and an ANOVA indicated that means were significantly different between regions for both species richness [$F(1,34)=17.38$, $p < 0.001$] and Shannon-Weiner index [$F(1,34)=10.05$, $p < 0.001$]. Post-hoc Tukey's HSD tests provided in

the Supplementary information (**TABLE 2.3** & **TABLE 2.4**) indicated perhaps surprisingly, that participants from Western Europe, the region with the poorest identification accuracy, provided the most rich and diverse answers (**FIGURE 2.3**).

TABLE 2.3 Post-hoc Tukey's HSD test displaying mean difference, 95% confidence intervals, and adjusted p-values for pairwise comparisons of richness of wrong answers from consumers among six European countries. The star (*) indicates significant pairwise comparison.

Pariwise comparison	Difference between means	Lower Confidence Interval	Upper Confidence Interval	Adjusted p-value	
Greece-Belgium	-5.500	-12.275	1.275	0.16556660	
Ireland-Belgium	5.000	-1.775	11.775	0.24798450	
Italy-Belgium	2.333	-4.441	9.108	0.89772450	
Spain-Belgium	-4.833	-11.608	1.941	0.28097880	
UK-Belgium	**8.167**	**1.392**	**14.941**	**0.01107001**	*
Ireland-Greece	**10.500**	**3.725**	**17.275**	**0.00068234**	*
Italy-Greece	**7.833**	**1.059**	**14.608**	**0.01611558**	*
Spain-Greece	0.667	-6.108	7.441	0.99964210	
UK-Greece	**13.667**	**6.892**	**20.441**	**0.00001330**	*
Italy-Ireland	-2.667	-9.441	4.108	0.83466790	
Spain-Ireland	**-9.833**	**-16.608**	**-3.059**	**0.00154388**	*
UK-Ireland	3.167	-3.608	9.941	0.71396190	
Spain-Italy	**-7.167**	**-13.941**	**-0.392**	**0.03327665**	*
UK-Italy	5.833	-0.941	12.608	0.12367760	
UK-Spain	**13.000**	**6.225**	**19.775**	**0.00003040**	*

TABLE 2.4 Post-hoc Tukey's HSD test displaying mean difference, 95% confidence intervals, and adjusted p-values for pairwise comparisons of diversity of wrong answers (as calculated with a Shannon diversity index) from consumers among six European countries. The star (*) indicates significant pairwise comparison.

Pariwise comparison	Difference between means	Lower Confidence Interval	Upper Confidence Interval	Adjusted p-value	
Greece-Belgium	-0.797	-1.746	0.151	0.14010026	
Ireland-Belgium	0.020	-0.929	0.969	0.99999984	
Italy-Belgium	-0.010	-0.959	0.939	0.99999999	
Spain-Belgium	-0.631	-1.580	0.318	0.35398819	
UK-Belgium	0.373	-0.575	1.322	0.83457765	
Ireland-Greece	0.817	-0.132	1.766	0.12352586	
Italy-Greece	0.787	-0.162	1.736	0.14931113	
Spain-Greece	0.166	-0.782	1.115	0.99430032	
UK-Greece	**1.171**	**0.222**	**2.119**	**0.00887008**	*
Italy-Ireland	-0.030	-0.979	0.919	0.99999870	
Spain-Ireland	-0.651	-1.599	0.298	0.32132617	
UK-Ireland	0.354	-0.595	1.302	0.86332097	
Spain-Italy	-0.621	-1.569	0.328	0.37150037	
UK-Italy	0.384	-0.565	1.332	0.81872116	
UK-Spain	**1.004**	**0.056**	**1.953**	**0.03310421**	*

A closer look at the diversity of answers in each of the two studied regions offers an interesting perspective on consumers knowledge and perceived fish diversity (**FIGURE 5.4.a & b**). Participants from WEU countries were more prone to take wild guesses, with some answers clearly illustrating a lack of knowledge regarding shape and morphology, such as 'goldfish', 'minnow', 'piranha', 'stickleback', and 'tiger shark' (**FIGURE 5.4.a**). Answers from SEU participants proved more realistic, with most wrong guesses attributed to other frequently consumed species (**FIGURE 5.4.b**). A species-by-species approach illustrated how these regions are accustomed to different fish species. For example, whereas sole was almost systematically correctly identified in southern countries, with only a few participants labelling it as a different species, more than half of the participants from WEU countries wrongly identified sole as plaice (*Pleuronectes platessa*) instead, which is the most popular flatfish traditionally caught and served in North-WEU waters (**FIGURE 5.4.a**). In contrast, participants from WEU countries accurately identified salmon more frequently than participants from SEU countries, who disproportionally guessed that

salmon and cod were a species of trout (**FIGURE 5.4.b**), most likely due to their habit of consuming imported cod and salmon in the form of fillets and/or salted and smoked products. Seabass, which only recently became a regular farmed import in North-western Europe, was mistaken for a greater variety of improbable species in WEU countries (**FIGURE 5.4.a**).

More efficient supply chains and urbanisation have dramatically modified seafood culinary cultures around the world (Bellotti, 2010; Carroll, 2009; Levin & Dufault, 2010). In 2018, over 70% of the European population lived in urban centers (i.e. cities or towns and suburbs) (Eurostat, 2020). The 'rural flight' phenomenon has, in part, been made possible by efficient modern supply chains, effectively disconnecting consumers from the source of the food they purchase (Bellotti, 2010). Today, produce does not need to be consumed fresh and can be shipped thousands of kilometers from its point of origin, sometimes traveling in convoluted ways before reaching its final destination (e.g. a large proportion of Atlantic cod are caught in Scandinavian waters and in the Barents sea, processed in China, only to find their way back to the European market) (Anderson et al., 2018). That being said, homogenisation of consumption habits is not entirely impermeable to cultural forces (Miller, Jessel, et al., 2012). Traditional staple foods which often used to be determined by local availabilities carry a strong momentum in guiding contemporary consumer choices (Asp, 1999; Nestle et al., 1998). Indeed, both WEU and SEU countries have had in the past and continue to have a significant supply of locally harvested fishery and aquaculture products (EUMOFA, 2017) which is highlighted in this study by region-specific abilities to identify commonly consumed fish.

Cultural habits have also led to a sustained demand for specific seafood products, yet with the decline of local fisheries, these demands have become harder to satisfy, leading to the commoditization of seafood, a phenomenon likely affecting consumers ability to discriminate between species. The whitefish market for instance, which was

dominated by cod for decades, also comprise other wild caught and farmed species (Asche et al., 2009) ever diluting an already broad category of species, and rendering the 'whitefish' term meaningless to many. Indeed, as indicated by Asche et al. (2009), the whitefish market also comprised other wild caught species such as saithe (*Pollachius virens*), redfish (*Sebastes* sp.), haddock (*Melanogrammus aeglefinus*), Pacific cod (*Gadus macrocephalus*), and Alaska pollock (*Gadus chalcogrammus*), most of which were deemed less valuable. The 1990s saw the introduction of widely traded farmed species in the whitefish market such as catfish (Siluriformes), hoki (*Macruronus novaezelandiae*), Nile perch (*Lates niloticus*), and more recently tilapia (*Oreochromis* sp.) and pangasius (*Pangasianodon hypophthalmus*) have dominated the market in terms of quantity (Asche et al., 2009). In a world where wild caught landings have plateaued and where aquaculture shoulders the growth of the seafood trade, staple species are increasingly becoming replaceable.

With urbanisation comes detachment from the natural world, which likely determines the results we observe. Providing consumers with adequate educational tools such as fish guides and certification schemes is useful for concerned and environmentally aware consumers; but when purchasing seafood, many consumers will often remain more focused on prices and expiry dates rather than on biodiversity (Pieniak et al., 2007). Knowledge, which is a decisive factor when it comes to the types of information consumers will consider while carrying out purchases (Pieniak et al., 2013), is highly heterogenous and will depend on factors such as education level, age, social status, culture, and rate of seafood consumption. For many less informed consumers "*fish is simply fish*" (Pieniak et al., 2007) and purchasing choices will be centered around quality and safety rather than around species type (Dey et al., 2008). The ability to identify wild species is linked with increased care, interest, and awareness for biodiversity (Mohneke et al., 2016; Schlegel & Rupf, 2010), and it should therefore come as a concern that the consumers from this study performed so poorly, with an average identification accuracy of ~30%. This comes in sharp contrast with performance on the identification of common

native terrestrial species, with an average identification accuracy of ~68% among the general public in the Netherlands (Hooykaas et al., 2019). To view fish products as a homogenised commodity is to overlook any species-specific ecological and environmental concerns that pertain to the diverse and complex fisheries industry. Until consumers understand this, seafood will likely continue to be at the mercy of fraudulent practices such as species mislabeling, species substitution, and illegal, unreported, and unregulated fishing. However, additional work is needed to evaluate if mislabeling practices and instances of seafood fraud tend to be higher in countries with lower seafood literacy.

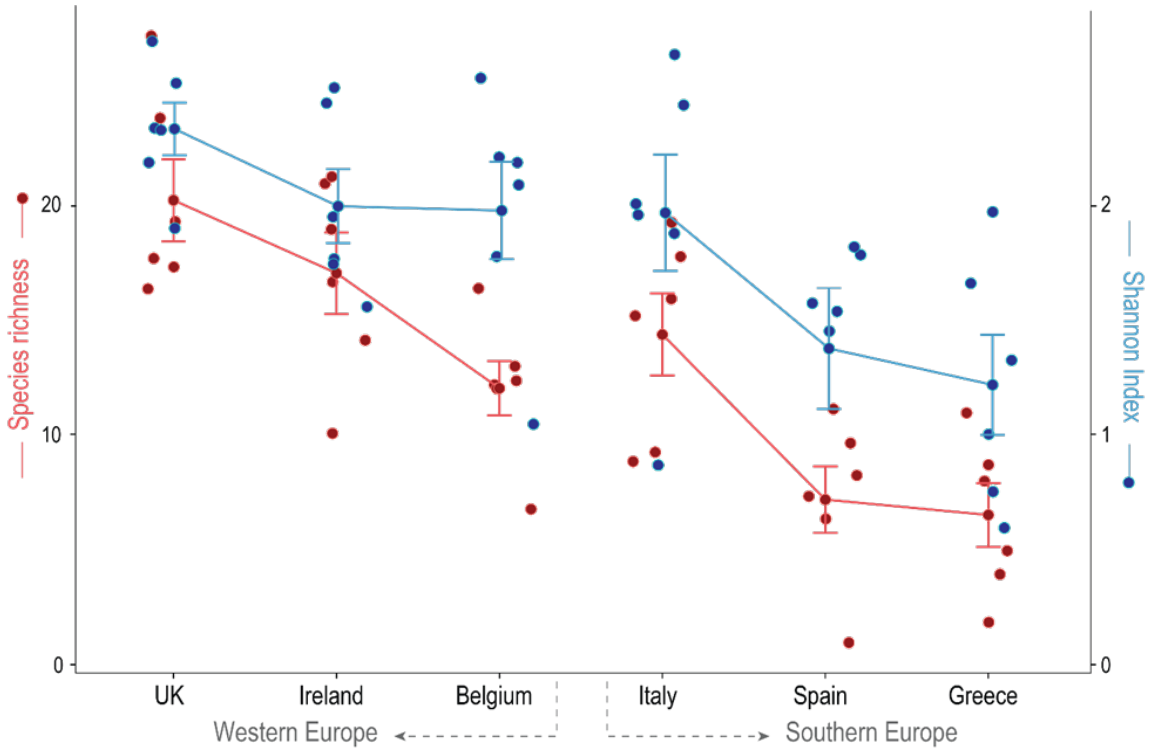
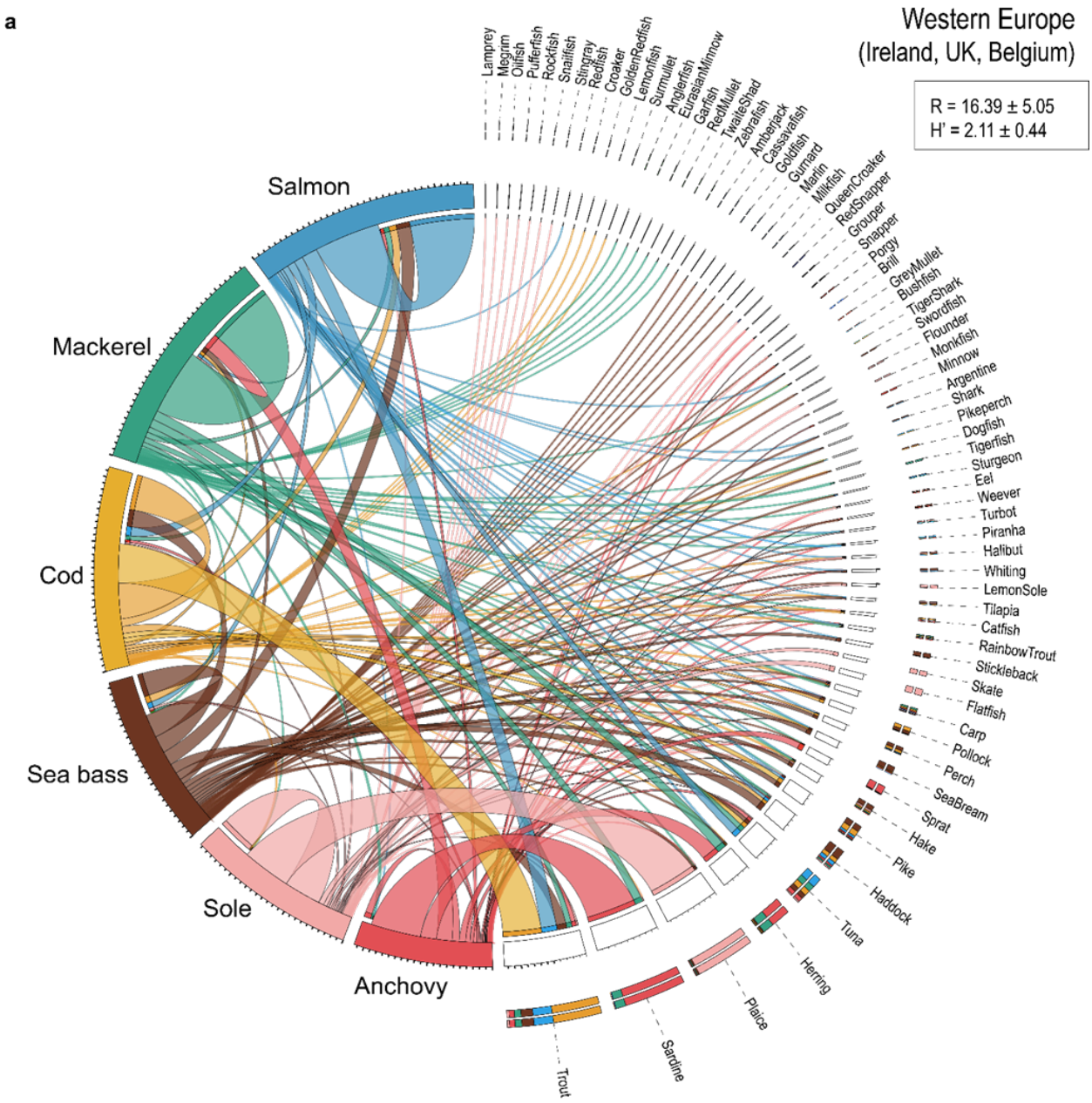


FIGURE 2.3 Line plots illustrating the mean values for species richness (in red) and Shannon diversity index (in blue) from the participants' wrong answers. An ANOVA indicated that means were significantly different between regions for both species richness ($F(1,34)=17.38$, $p=1.99\times10^{-4}$) and Shannon-Weiner index ($F(1,34)=10.05$, $p=3.2\times10^{-3}$).


a
Western Europe
(Ireland, UK, Belgium)
R = 16.39 ± 5.05
H' = 2.11 ± 0.44
Salmon
Mackerel
Cod
Sea bass
Sole
Anchovy
Lamprey
Megrim
Oilfish
Pufferfish
Rockfish
Snailfish
Stingray
Redfish
Croaker
GoldenRedfish
Lemonfish
Surmullet
Anglerfish
EurasianMinnow
Garfish
RedMullet
TwaiteShad
Zebrafish
Amberjack
Cassavafish
Goldfish
Gurnard
Marlin
Milkfish
QueenCroaker
RedSnapper
Grouper
Snapper
Porgy
Brill
GreyMullet
Bushfish
TigerShark
Swordfish
Flounder
Monkfish
Minnow
Argentine
Shark
Pikeperch
Dogfish
Tigerfish
Sturgeon
Eel
Weever
Turbot
Piranha
Halibut
Whiting
LemonSole
Tilapia
Catfish
RainbowTrout
Stickleback
Skate
Flatfish
Carp
Pollock
Perch
SeaBream
Sprat
Hake
Pike
Haddock
Tuna
Herring
Plaice
Sardine
Trout

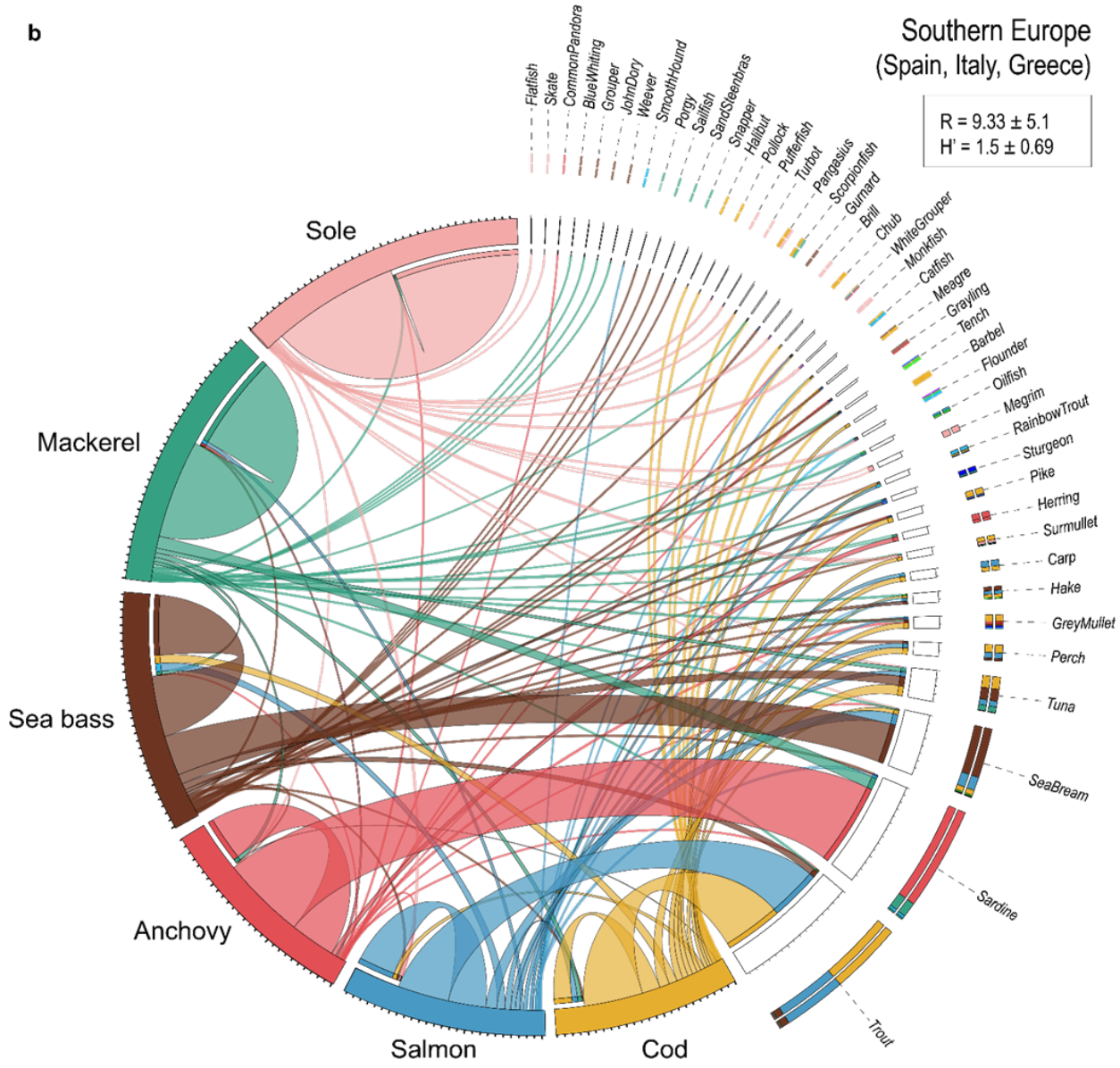


FIGURE 2.4 Diversity of guessed answers by consumers linked to the six species presented to customers in the form of pictures (to the right of the diagram = sole, mackerel, sea bass, anchovy, salmon, and cod). The bands thickness from the circular diagrams indicate the proportions of answers linked to a given guess. Answers were cumulated for **(a)** Western European countries (Ireland, United Kingdom, and Belgium) and **(b)** Southern European countries (Spain, Italy, Greece). Mean species richness and mean Shannon-Weiner diversity was calculated for each region. An ANOVA indicated that means were significantly different between regions for both species richness ($F(1,34)=17.38$, $p=1.99\text{x}10\text{-}4$) and Shannon-Weiner index ($F(1,34)=10.05$, $p=3.2\text{x}10\text{-}3$). The circular diagram was created using the online Circos software (Krzywinski et al., 2009).

2.5 Future Directions

Asking people to identify staple fish species is only one of the many ways to explore seafood literacy. The results highlighted in this study illustrate the need to investigate further the question of consumers familiarity with the seafood they consume. Interrogating consumers on their reasons for purchasing or not a given seafood would help better understand some of the motivations that guide consumers in their purchases. Certain socioeconomic variables such as age, education level, income, dietary preferences, health conditions might be tightly linked with seafood consumption habits and could offer an important insight on some of the factors that affect seafood literacy. The study design with regards to the type of photography shown or location sampled could also be explored further. For example, most people are exposed to fish filets rather than whole fish (as might be the case with salmon); showing fish filets might therefore give drastically different answers. Furthermore, some of the most consumed species such as tuna were not included in this study and would need to be added in future studies. The location of the sampling is also likely to affect the response accuracy with coastal communities possibly performing better and inland rural residents not performing as well. An important point to consider is the number of replicates per countries which would allow to extrapolate the results to entire nations with higher confidence.

2.6 Conclusions

We provide the first quantitative characterisation of seafood literacy in the largest seafood market in the world, the European Union. Though consumers from urban centers of Southern European countries were generally more successful at identifying fish, the overall trend from this study illustrates consumers' relative unfamiliarity with the appearance of commercial fish species in both Western and Southern European countries. Culinary traditions heavily influence the diversity of answers from the respondents, and the correct identification of species. Nevertheless, despite cultural

inertia, modern lifestyles are increasingly detaching consumers from the source of their food. The growing demand for seafood and the dynamic nature of the fisheries industry are responsible for the grouping of fish species into broad categories, masking seasonal and yearly variations in species harvest, production, and trade, and offering a constant and steady commodity to the consumer. Globalised supply chains and urbanisation are likely at the core of the low seafood literacy we observe, and of a gradual homogenisation of food cultures among countries. In this context, the low identification accuracy we observe across countries seems predictable, if anything, it is rather surprising that consumers should be able to identify any fish species at all, and it reveals some level of contact with whole specimens despite market trends pushing for processed products.

Increasingly, environmental Non-Governmental Organisations and stakeholders have provided consumers with seafood educational support as an attempt to guide their purchase towards sustainable consumption habits. Despite well-intentioned incentives provided by labelling standards and certification schemes, our results clearly demonstrate a lack of knowledge regarding the appearance of commonly consumed fish. Studies have demonstrated that the ability to identify given species is linked with increased appreciation and respect and can have far reaching consequences in terms of conservation. Though this study does not examine causality between seafood literacy and seafood sustainability, it is nonetheless possible to assume that consumers inability to identify common fish species might result in a disinterest for sustainable consumption of seafood. We therefore suggest that there is a pressing need to continue educating consumers about seafood and we recognize that this responsibility should not solely rest on the shoulders of NGOs. It ought to be addressed by concerted actions involving practicing scientists, governments, and civil society at large, with the view to transform the way modern citizens perceive wild foods. It is not surprising that in the current context, consumers might display a lack of interest and ability in discriminating between species. When it comes to identifying fish species, consumers are like fish out of water.

Chapter 3

Field application of portable, universal real-time DNA barcoding for reliable fish species authentication

Authors: **Cusa, Marine**; Sanchez, Aquiles; Goodhead, Ian; Mariani, Stefano

3.1 Abstract

DNA barcoding has become a mainstay of fish species identification applications worldwide. The widespread phenomenon of seafood mislabelling poses a risk to consumers' health, economic growth, and resource and conservation management. Despite the success of DNA barcoding in improving seafood traceability, leading to greater public awareness and more stringent regulations in many parts of the world, there remains a need for simpler, rapid, portable, and universal species DNA authentication tools. FASTFISH-ID™ is an innovative PCR method of closed-tube DNA barcoding that offers a promising rapid and cost-effective solution for fish species authentication. The FASTFISH-ID™ method uses a single qPCR assay targeting segments of the mitochondrial COI barcode region, generating species-specific melt curves (fluorescent signatures) that can be compared against databases of existing melt-curve profiles for rapid species identification. To evaluate the reliability and robustness of this method in an in-field setting, we tested FASTFISH-ID™ using a portable real-time PCR instrument aboard the Research Vessel CEFAS Endeavour during a three-week survey off the South-West coast of England and in the English Channel, using a random set of commercially important pelagic ichthyofauna collected with a mid-water trawl. FASTFISH-ID™ rapidly generated unique species-specific fluorescent signatures for 14 out of the 15 species which were sampled opportunistically aboard the trawler, including when operating in rough weather. Additionally, the method also allowed for haplotype discrimination of European anchovy. These findings not only demonstrate that the method performs robustly in the field, but also that it has potential for discriminating population units within a species. The resulting species-specific fluorescent signatures independently validated by DNA sequencing will serve as the basis to construct a reference database for identification of North Atlantic species. This study is the first to validate FASTFISH-ID™ for field use, opening a wide panel of new applications for this method.

Keywords

Species identification · Field tool · On-site · qPCR · Seafood fraud · Seafood mislabelling

3.2 Introduction

3.2.1 Seafood mislabelling and forensic authentication tools

The growing demand for seafood is leading to an ever-increasing pressure on marine ecosystems. According to the latest Food and Agriculture Organization (FAO) report on the state of the world fisheries and aquaculture (FAO, 2020), 2018 was unprecedented in terms of fish production, reaching 178 million tons as a result of both fisheries and aquaculture. The same report shows that total seafood consumption rose by 122% since 1990. About half of this production comes from wild capture fisheries, making sustainable exploitation a key element to meeting demand and preserving quotas. Both resource management and governance are crucial if wild seafood resources are to continue being exploited at this rate, but success heavily relies on compliance and appropriate enforcement. The increasing demand for seafood products is leading to a greater diversity of species being traded in an ever more complex supply chain (Cawthorn et al., 2018; Di Muri et al., 2018) and inexorably opens the door to malpractice (Sumaila et al., 2020).

Applied wildlife forensics has benefited greatly from recent advances in genetics (Gouda et al., 2020), and deoxyribonucleic acid (DNA)-based tools have uncovered a concerning phenomenon present in the seafood industry: the mislabelling of retailed seafood products. A search on *pubmed.gov* revealed that over the last decade (2010-2020 = 168 publications), papers including the words “seafood” plus one of “substitution”, “authenticity” or “mislabelling” have increased by a factor of 7 compared to the previous decade (2000-2010 = 25 publications) (**FIGURE 3.1**). The substitution of species in the seafood industry is a widespread phenomenon that continues to be denounced at an unprecedented scale, with annual increases in global publications exposing the issue (Luque & Donlan, 2019). The matter has not only been noted in the peer reviewed literature (Donlan & Luque, 2019; Fox et al., 2018; Luque & Donlan, 2019) but also in reports from Non-Governmental Organizations (NGOs) (OCEANA, 2016) and government bodies (FAO, 2018a), and it is regularly discussed in the general media (Barrie, 2018; Doward, 2017; Leahy, 2021; Mariani et al., 2014). The mislabelling rate itself is hard to predict as it varies widely between countries (Silva & Hellberg, 2021) and more so with

certain groups of species, with the Lutjanidae ('snappers') family appearing to have some of the highest incidents of mislabelling (>50%) among all studied species (Cawthorn et al., 2018; Luque & Donlan, 2019). Mislabelling facilitates products from Illegal, Unreported and Unregulated (IUU) fishing or poorly managed stocks to enter the market (Kroetz et al., 2020; Sumaila et al., 2020) and find their way into the plates of consumers who, despite mindful incentives, are unwillingly perpetuating the issue by purchasing such products. The practice can be accidental or deliberate and happens at any node along the supply chain (Donlan & Luque, 2019; Fox et al., 2018) although it is most often reported at the retail level (supermarkets, fishmongers, and restaurants) (Miller, Jessel, et al., 2012; Shehata et al., 2019). It represents a major challenge in the seafood industry and must be tackled urgently if instances of fraud are to be controlled and if sustainable fishing is to be prioritized.

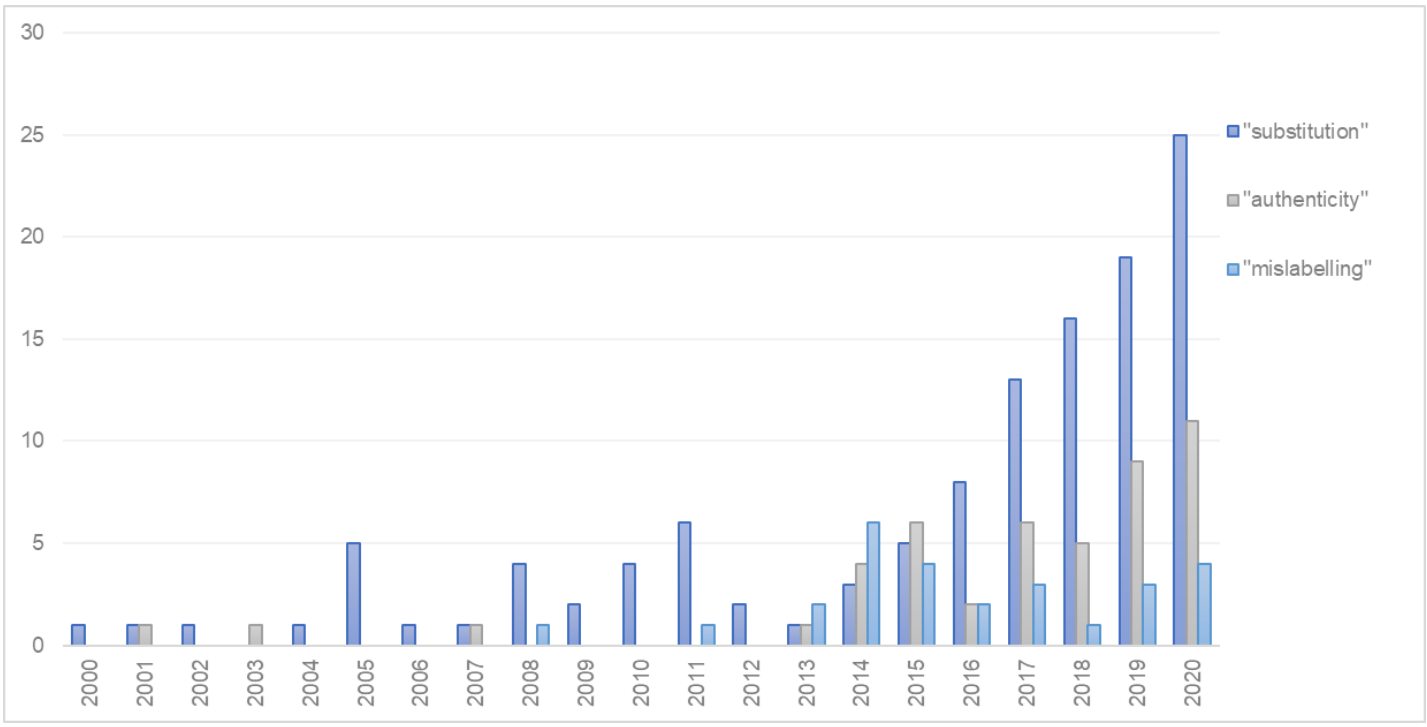


FIGURE 3.1 Number of publications from the year 2000 to 2020 containing the word seafood plus substitution and/or authenticity and/or mislabelling from the website pubmed.gov.

DNA-based methods can be powerful tools in identifying mislabelled marketed specimens and they have proved to be the forensic tools of choice among the majority (about 94%) of peer-reviewed work focusing on that topic (Luque & Donlan, 2019). These methods have offered a platform for the identification of marketed seafood products at any stage of the supply chain (Fox et al., 2018) which has led to the prosecution of entities conducting or supporting fraudulent fishing practices (Martinsohn et al., 2019). Due to its ubiquitous nature, high inter-specific species variation, and low degradation rate, DNA is a particular useful marker when attempting to identify whole or processed fish, even in conditions of heavy processing (canned, dried, salted, cooked, etc.) (Silva & Hellberg, 2021). Sanger sequencing of the cytochrome C oxidase I (COI) gene found on the mitochondrial DNA of animal species has established itself as the seafood forensic method of choice (Fernandes et al., 2020). Other methods are continuously being explored, however, as the demand for easier, rapid authentication tools continues to rise. Species-specific PCR assays have offered a rapid and potentially portable alternative and have proven quite successful in the identification of seafood samples and in the seizing of fraudulently traded seafood commodities (Caballero et al., 2012; Cardeñosa et al., 2018, 2019), but they remain limited by the need to develop unique species-specific assays. TaqMan probes which are often used in species-specific real-time PCRs are designed to bind to a target sequence and emit fluorescence when a quencher is separated from the fluorophore during the amplification process. High Resolution Melting analysis (HRM) which is based on the denaturation of DNA at different temperatures depending on the type of nucleotide bonds (GC bonds requiring slightly higher temperatures to denature than AT bonds) allows for the generation of unique melt curves when sequences differ even by a single base pair. This is developed as a tool in the identification of closely related species such as salmonids (Shi et al., 2020) or *Mytilus* spp. (Quintrel et al., 2021). In contrast, the FASTFISH-ID method uses mismatch tolerant fluorescent probes that bind on single strands of DNA at the end of the PCR cycle and emit fluorescent signals as they denature during a final PCR step where the temperature is gradually increased. More recently, third generation sequencing such as nanopore technologies have largely demonstrated their utility and reliability as portable field sequencing devices (Menegon et

al., 2017; Pomerantz et al., 2018), and are rapidly gaining ground in the domain of seafood authentication (Ho et al., 2020; Johri et al., 2019) though they continue to be an expensive technology compared to conventional or real-time PCR.

Despite research initiatives to develop DNA-based Standard Operating Procedures (SOPs) (FishPopTrace, 2013; Labelfish Consortium, 2014), governmental institutes lag behind predominantly due to a misunderstanding of the capabilities, cost, and performance of genetic tools (Bernatchez et al., 2017; Martinsohn et al., 2019). These arguments seem hardly justifiable in light of the collective endeavour to communicate on cutting-edge technologies and to assemble robust and reliable SOPs. However, other arguments such as the lack of accessibility to genetic tools and a desire for cheaper methods could be addressed by offering methods that are indeed cheaper, faster, portable, and more readily available to non-experts. The ideal tool would allow fisheries scientists, managers, and the industry itself to conduct *in situ* analysis without relying on external expertise and laboratories (i.e. for sequencing). The use of a species-specific quantitative Polymerase Chain Reaction (qPCR) eel assay in a real enforcement scenario leading to prosecution (Cardeñosa et al., 2019) showcases that if quick, easy, portable, and reliable DNA-based tools are made available to authorities, they can and will be used in the context of enforcement. qPCR assays have proved particularly promising due to their simplicity, rapidity, portability, reproducibility, and cost, but remain limited by the need to develop species-specific probes and assays.

3.2.2 Closed-Tube barcoding: a promising tool for seafood authentication

Recently, Closed-Tube DNA barcoding was advanced as a promising solution with the potential to act as a portable and universal seafood authentication method (Naaum et al., 2021 - See Appendix). Known as FASTFISH-ID™, this method bypasses the need for sequencing and offers a potentially reliable, inexpensive, rapid, and universal seafood species authentication tool. It combines the non-targeted species identification afforded

by DNA barcoding with the speed and portability of real-time PCR tests. Closed-Tube DNA barcoding uses asymmetric PCR amplification (Sanchez et al,. 2004; Pierce et al., 2005) and a single set of specially designed consensus probes (Rice et al., 2012) to amplify and analyze two specific regions within the COI segment, barcode segment 1 (BS1) and barcode segment 2 (BS2) regions that exhibit high levels of inter-species sequence variation (Naaum et al, 2021). The method converts species-specific BS1 and BS2 sequences into unique fluorescent signatures. Species are then defined by the unique combinations of BS1 and BS2 signatures. For species identification, the BS1 and BS2 fluorescent signature set of a specimen is compared against a cloud-based reference library of species-specific fluorescent signatures generated from vouchered specimens. The entire process can be carried out on a portable thermocycler, such as the Biomolecular Systems MIC, which is capable of examining 48 samples at a time and results are generated in about two hours, in a single-tube format.

The universal probe set used for FASTFISH-ID™ was designed via *in silico* sequence testing and statistical analysis of BS1 and BS2 regions from large numbers of COI target sequences selected from 200 commercial fish species (Naaum et al., 2021). Although the method has been validated in a controlled laboratory environment for 18 commercially important fish species, its in-field application and universality for species identification – i.e. the extent to which FASTFISH-ID™ can identify any random fish species of commercial importance – remain untested. Furthermore, the idea of targeting short fragments of the COI barcode with fluorescent probes to generate species-specific fluorescent curves is ingenious but could prove problematic in the case of intra-specific variations due to Single Nucleotide Polymorphisms (SNPs) in the target regions, or heteroplasmy (the presence of more than one type of mtDNA genome). Considering how valuable such method could be to the seafood sector in terms of enforcement and authentication, we investigated the suitability of the closed-tube barcoding technology for rapid and accurate on-site determination of pelagic fish species caught via mid-water trawl onboard a fisheries survey vessel bound for the Celtic Sea and the English Channel.

3.2.3 Objectives

To gauge whether this method is fully functional in the field and evaluate its universality and reliability we i) observed whether unique and robust fluorescent signatures were successfully generated on-site in a variety of scenarios; ii) compared the quality of the fluorescent signatures generated in the field with signatures generated in a controlled laboratory environment; iii) verified if field specimens could be correctly identified once voucher specimens were added to the reference library and evaluated how species variants might affect results reliability, and iv) explored the use of the method for identification of European anchovy haplotypes.

3.3 Methods

3.3.1 Evaluating the practicality of FASTFISH-ID™ in the field

Portable Mic PCR laboratory setup

The main devices used for running the FASTFISH-ID™ protocol in the field were the Bio-Molecular Systems Mic PCR cycler, a microfuge (Personal Microcentrifuge, Cole-Parmer, Vernon Hills, IL, US), vortex mixer (Vornado Miniature. Benchmark Scientific, Edison, NJ, US) and a Windows 10 laptop. The requirements for using the FASTFISH-ID™ method in the field were rather minimal and could be easily met aboard the RV CEFAS Endeavour: (1) a 230V plug in order to add a multi-plug to power the various electronic devices and (2) a small bench space where the DNA extraction and PCR amplification in the Mic could be conducted continuously over the span of 4 to 5 hours (**FIGURE 3.2**). Any heating required for the DNA extraction was performed using the Mic qPCR device itself. The reagents have been enhanced to be stable at room temperature for up to three months and do not require a freezer for storage during field work.

Sampling and tissue processing

Pelagic trawling was conducted almost every single day, sometimes several times a day depending on the presence and density of an acoustic backscatter. The trawl content was first processed by CEFAS specialists who sacrificed the specimen and identified them taxonomically down to species level. Once the fish were processed and identified, specimens were then selected for the analysis. We randomly sampled up to 25 specimens for each species and took a picture of each specimen along with an ID number (**FIGURE 3.2**). If possible, the specimen selected for the analysis were caught in different trawls to maximize intra-specific species variation but not all species were caught more than once. We then collected a fin clip of roughly one cm long, or the whole pectoral fin for small specimens, and placed them in a tube with 95% Ethanol with their associated ID number. Tweezers, scalpels, and dissecting scissors were cleaned with DNA-away Surface Decontaminant in between each sample to avoid contamination.

Field DNA extraction and Mic PCR product amplification

Up to five individuals per species were processed for DNA extraction in the field. The tissue of a given sample was taken from the storage tube and excess ethanol was removed by dabbing the tissue on clean paper towel. A 2 mm clip was then cut and placed in a Mic PCR tube. 100 µL potassium hydroxide (KOH) lysis buffer (Tagliavia et al. 2016) was added to the tube, and the tubes were incubated at 85°C for 15 minutes using the Mic. The lysate from the Mic PCR tube was then transferred to another tube containing 300 µL of 100 mM Tris-HCl pH 8.3 (modified after Tagliavia et al. (2016). After thorough mixing using the mini vortex, the DNA was ready for use or stored in the fridge at 4°C. FASTFISH-ID™ qPCR amplification was done according to Naaum et al. (2021). Briefly, ThermaGenix PCR reagents (ThermaGenix, Natick, MA) containing PCR buffer, MgCl2, dNTPs, FDA degenerate primers, FASTFISH-ID probe mix, ThermaStop, Cal-Red, 610, Quasar 670, ThermaMark, and MyTaq DNA polymerase were added to extracted fish DNA (2 µL) for a final volume of 12.5 µL.

The PCR profile consisted of 94 ◦C for 2 min, followed by 5 cycles of 94 ◦C for 5 sec, 55 ◦C for 20 sec, 72 ◦C for 45 sec, followed by 65 cycles of 94 ◦C for 5 sec, 70 ◦C for 45 sec. To enable probe annealing, the temperature was then lowered to 40 ◦C for 10 min, and the temperature was then increased from 40 ◦C to 87 ◦C at 0.1 ◦C/sec for fluorescent acquisition thereby creating melting curves which are translated into fluorescent signatures (the first derivative of the resulting melting curves).

Fluorescent signature interpretation

FASTFISH-ID™ identifies species by comparing the fluorescent signatures from a tested specimen against a reference library of species-specific fluorescent signatures prepared from vouchered specimens. The fluorescent signatures obtained on the boat could be compared to each other visually to determine whether they were species-specific but could not immediately be scrutinized using the FASTFISH-ID™ library as the reference library did not contain the randomly selected species at the time of the field validation, except for Haddock (*Melanogrammus aeglefinus*). Consequently, a first examination of the signatures was performed visually onboard, and a more thorough species authentication was performed *post hoc*, once additional voucher specimens had been added to the online reference library.

3.3.2 Evaluating the universality and reliability of the FASTFISH-ID™ technology

Sample processing in a controlled laboratory environment

The fluorescent signatures obtained on board the R/V CEFAS Endeavour were compared to those obtained in a controlled laboratory environment from a different set of individuals of the same species. Depending on the number of specimens collected, from one to 17 samples were selected for further analysis in a controlled laboratory environment for the following ten species: *Dicentrarchus labrax*, *Sardina pilchardus*, *Scomber scombrus*,

Trisopterus esmarkii, *Merlangius merlangus*, *Micromesistius poutassou*, *Clupea harengus*, *Engraulis encrasicolus*, *Thunnus thynnus*, and *Merluccius merluccius*. DNA extraction of these samples was performed on land on 30 mg of fin sample with the E.Z.N.A. Tissue DNA Kit (Omega Bio-tek, Inc. Norcross, GA, USA). The samples were then processed using the two workflows for species identification illustrated in **FIGURE 3.3**. For the standard symmetric PCR amplification, the samples were processed following the protocol from the Labelfish SOP (Labelfish Consortium, 2014). PCR products were prepared using the ready-to-use Qiagen multiplex PCR kit (Qiagen, Germantown, MD, USA) with a cocktail of four primers, VF2_t1, FishF2_t1, FishR2_t1, as described by Ward et al. (2005) and the FR1d_t1 primer described by Ivanova et al. (2007) to amplify the entire COI 650 base-pairs barcode region. PCR products were inspected by gel electrophoresis prior to out-sourced Sanger sequencing (Macrogen Inc., Amsterdam, The Netherlands). The same samples were also run in the Mic using the closed-tube barcoding protocol described in Naaum et al. (2021).

Data processing

The integrity of the DNA sequences was checked using the sangeranalyseR (Chao et al., 2020) and the sangerseqR (Hill et al., 2014) packages in R (R Core Team, 2019). These packages were used to trim low-quality bases from the reads, to trim the primers and to call and visualize secondary peaks in the sequence. The maximum base pairs per row was set to 100 and the height of each row in chromatograms was set to 200. The ratio of a secondary peak to primary peak in the chromatogram was set to 0.33. The resulting FASTA files were then imported into the Barcode of Life Database (BOLD; http://www.barcodinglife.org/) for species identification by querying the Species Level Barcode Records database. The BS1 and/or BS2 in the generated sequences were aligned with the program Aliview (Larsson, 2014) to identify SNPs present in either one or both of the FASTFISH-ID™ barcode segment regions.

Some species had to be removed from the analysis for the following reasons: insufficient specimens to generate robust fluorescent signatures (*T. thynnus*), DNA amplification failed under the controlled laboratory conditions as evidenced by the gel eletrophoresis results (*C. harengus* and *D. labrax*), no species match was found upon analysis of Sanger sequences (*D. labrax*), or too many variants existed for a single species (*E. encrasicolus*). As a result, out of the ten species originally selected, only the following six species were retained for further analysis: *S. pilchardus*, *S. scombrus*, *T. esmarkii*, *M. merlangus*, *M. poutassou*, and *M. merluccius*.

To determine whether the fluorescent signatures generated in the laboratory from the new species were truly unique among all the FASTFISH-ID™ species-specific fluorescent signatures obtained to date the new fluorescent signatures were first queried against the original FASTFISH-ID™ reference database. The new fluorescent signatures were then added to the reference database and queried against themselves to ensure that the online algorithm recognized each uniquely. Finally, the results generated on the boat were evaluated against this updated reference database to assess the success with which species randomly sampled and processed on-site would be correctly authenticated. The fluorescent signatures for the specimens that were not recognized were then compared to the ones in the library to understand why they failed to be distinguished. These reference library queries were performed using the FASTFISH-ID™ online scoring algorithm at https://thermagenix-speciesid.shinyapps.io/species-id_fastfish-id_answers_v2/ .

We determined the frequency of the BS1 and BS2 DNA sequence variants among all COI sequence information mined from the Barcode of Life (BOLD) database for a given species to establish if the fluorescent signatures we generated were representative of the majority of the specimens. Towards this goal, we downloaded from BOLD all known COI DNA barcode sequences for each chosen species as of July 2021, aligned those sequences to determine the location of the BS1 and BS2 segments and measured the frequency of the BS1 and BS2 DNA sequence corresponding to the generated fluorescent signature among the total number of intraspecies sequence variants for BS1 and BS2. This analysis also allowed us to predict the total number of possible fluorescent signatures

from a given species based on the total number of currently known BS1 and BS2 intra-species polymorphisms.

3.3.3 Exploring the use of closed-tube barcoding for intraspecies discrimination

The anchovy samples from this study displayed a variety of signatures and sequence variants, and their COI barcode sequence was scrutinized in AliView against results from Pappalardo et al. (2015) who identified six diagnostic SNPs within the European anchovy COI barcode region to discriminate among common and well studied anchovy haplotypes. We evaluated where these SNPs were located with respect to the binding regions of the fluorescent probes for fish species identification and explored whether the fluorescent signatures differed between haplotypes.

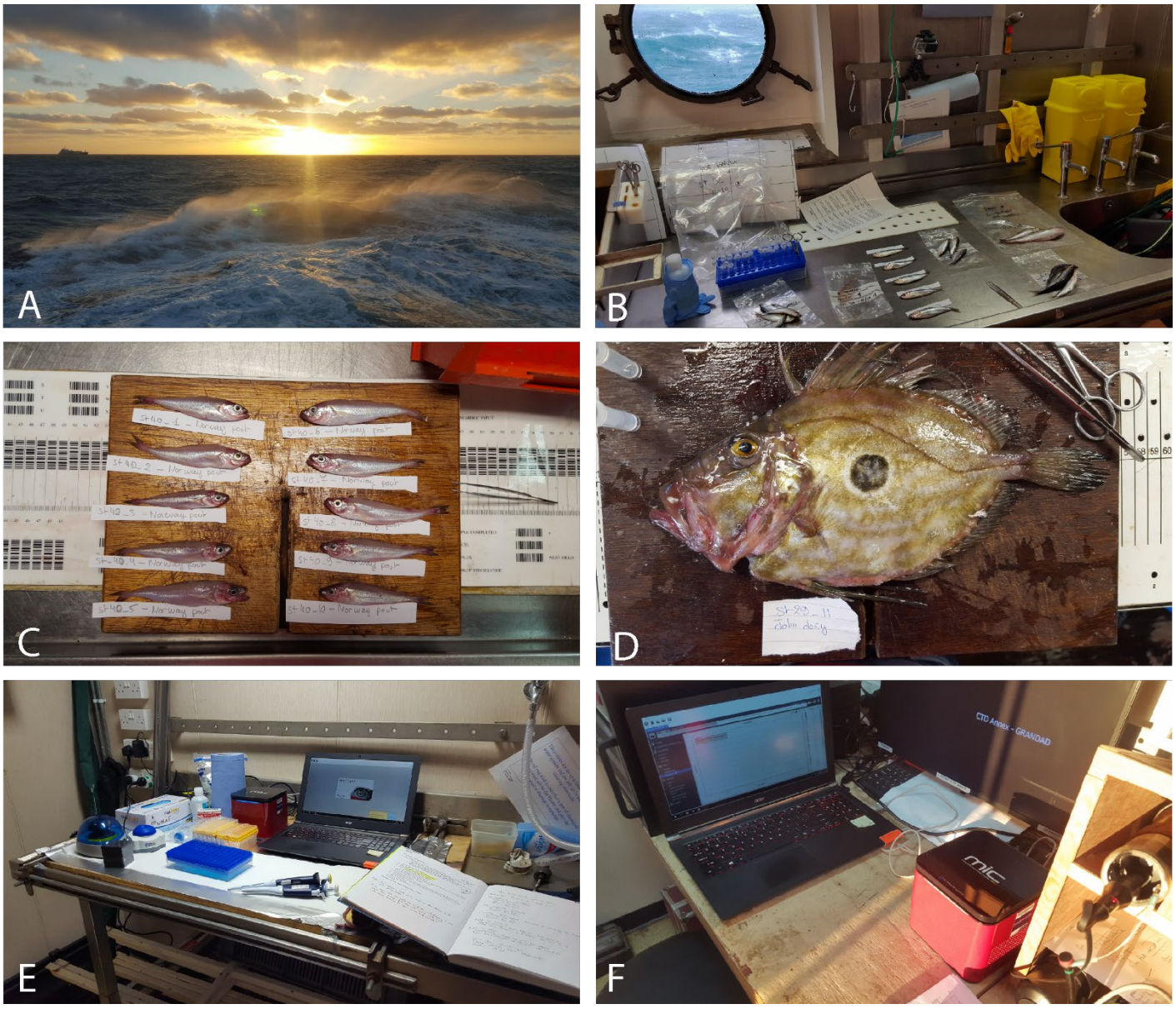


FIGURE 3.2 Illustration of the first field validation of the FASTFISH-ID™ method using a Magnetic Induction Cycler technology; (A) Field conditions in which the test was conducted, at sea aboard the RV CEFAS Endeavour; (B) A small workstation needed to be set up to take pictures and tissue samples from the fish that were opportunistically sampled from the pelagic trawls; Fish (C, Norway pout) and (D, John Dory) were visually identified and displayed on a wooden bench for pictures, they were assigned an ID number, and processed; (E) The set up for conducting the FASTFISH-ID™ method in one of the walk-in fridges (only room available) of the ship; (F) The Mic running the FASTFISH-ID™ profile, the fluorescent signatures starting to form on the connected laptop.

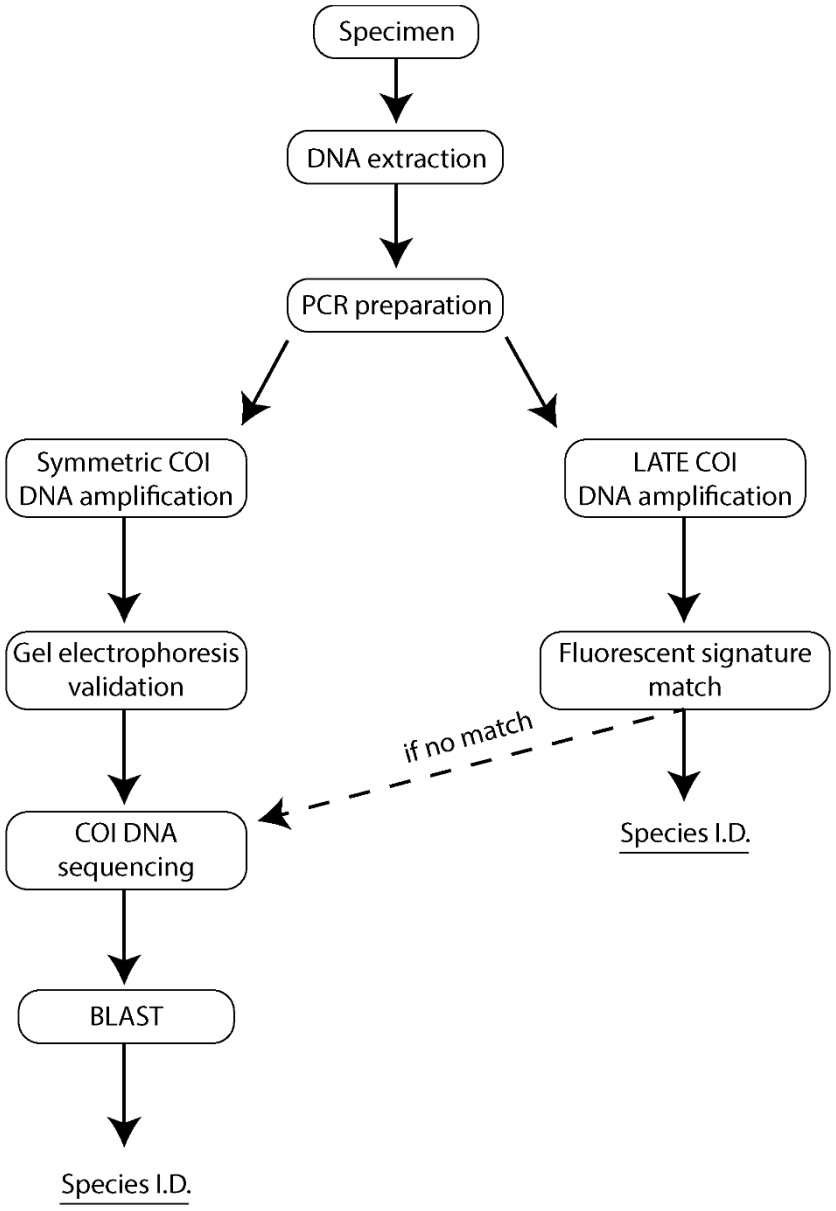


FIGURE 3.3 Schematic diagram illustrating the workflow for species identification using DNA barcode sequencing (left) and the FASTFISH-ID™ method (right). "LATE" stands for LATE-PCR (Linear After The Exponent PCR).

3.4 Results

3.4.1 Evaluating the practicality of FASTFISH-ID™ in the field

Field work conditions and portability of the FASTFISH-ID™ method

We sampled and categorized a total of 272 specimens that were then brought back to the University of Salford for further analysis. Out of these, 52 individuals from 15 species were processed on-board the R/V CEFAS Endeavour using the FASTFISH-ID™ protocol and reagents. Field conditions were quite variable, and the Mic ran successfully even while the ship was moving substantially due to rough weather. Pelagic trawling occurred almost every day and we opportunistically collected fish from these trawls. From the time the fish was brought back on deck to the time we generated a species-specific signature, a minimum of four hours elapsed, although this was also variable, and depended on weather conditions and other work obligations. On one occasion and due to poor weather forecast, we were able to extract the DNA and prepare the samples for a full Mic run in less than an hour, the run itself lasts about two hours. We could theoretically have processed 48 specimens in a single Mic run, but due to the validation nature of this study, we decided to test samples in triplicates, leaving room for testing only 15 individuals maximum and one control per run.

Out of a total of 52 specimens from 15 species randomly sampled aboard the R/V CEFAS Endeavour and processed using the FASTFISH-ID™ protocol, 48 specimens from 13 species generated strong and unique BS1 and/or BS2 fluorescent signatures, three specimens of European seabass (*D. labrax*) generated a weak signature, and the single specimen of Atlantic bluefin tuna (*T. thynnus*) caught and sampled failed to generate any signature possibly as a result of failed DNA amplification. More specifically, European sprat (*S. sprattus*), European pilchard (*S. pilchardus*), Hake (*M. merluccius*), and Mueller's pearlside (*M. muelleri*) displayed strong and unique signatures both for BS1 and BS2. Possible variants were identified for Norway pout (*T. esmarkii*), Blue whiting (*M. poutassou*), and Whiting (*M. merlangus*). Consistent and unique signatures were observed on BS2 for European anchovy (*E. encrasicolus*), Spiny dogfish (*S. acanthias*), Atlantic

mackerel (*S. scombrus*), and Atlantic herring (*C. harengus*) whereas John dory (*Z. faber*) displayed a strong and unique signature on BS1. The single Haddock caught and sampled was the only species tested that was in the FASTFISH-ID™ online reference database and was immediately and successfully matched back to *M. aeglefinus*.

Out of the 15 species analyzed on board, 10 were then analyzed using a different set of specimens in a controlled laboratory environment back on land. Those samples were processed using the FASTFISH-ID™ protocol and were subsequently sent for Sanger sequencing. In visually comparing the fluorescent signatures generated aboard the R/V CEFAS Endeavour, with those generated in a controlled laboratory environment, we observed that the signatures were comparable in strength (amplitude) and quality (smoothness). The unique species-specific signatures from the laboratory tested specimens matched those from the field, though this was then tested more robustly by using an updated version of the FASTFISH-ID™ database and considering possible variants that may affect authentication.

3.4.2 Evaluating the universality and reliability of the FASTFISH-ID™ technology

Laboratory validation

To evaluate the universality of the FASTFISH-ID™ method for species identification, we selected a total of 43 samples from 10 different species for comparative analysis under controlled laboratory conditions. Four out of these ten species were removed from further analysis due to failed DNA barcode amplification, lack of or poor species match in BOLD, or excessive number of variants for a single species. After confirmation of species identity via Sanger sequencing, the following samples were used to populate the updated FASTFISH-ID™ database: three specimens of Norway pout, three specimens of whiting, three specimens of blue whiting, three specimens of hake, seven specimens of European pilchard, and two specimens of Atlantic mackerel (**FIGURE 3.4**).

In an initial assessment where the fluorescent signatures from the selected species were queried against the original reference database, no matches were found, illustrating that their fluorescent signatures were unique with respect to those existing in the original database. When the updated database including the six new species was used, all species used to populate the updated database and tested in a controlled laboratory environment were correctly matched back, thereby illustrating that they each displayed unique species-specific fluorescent signatures.

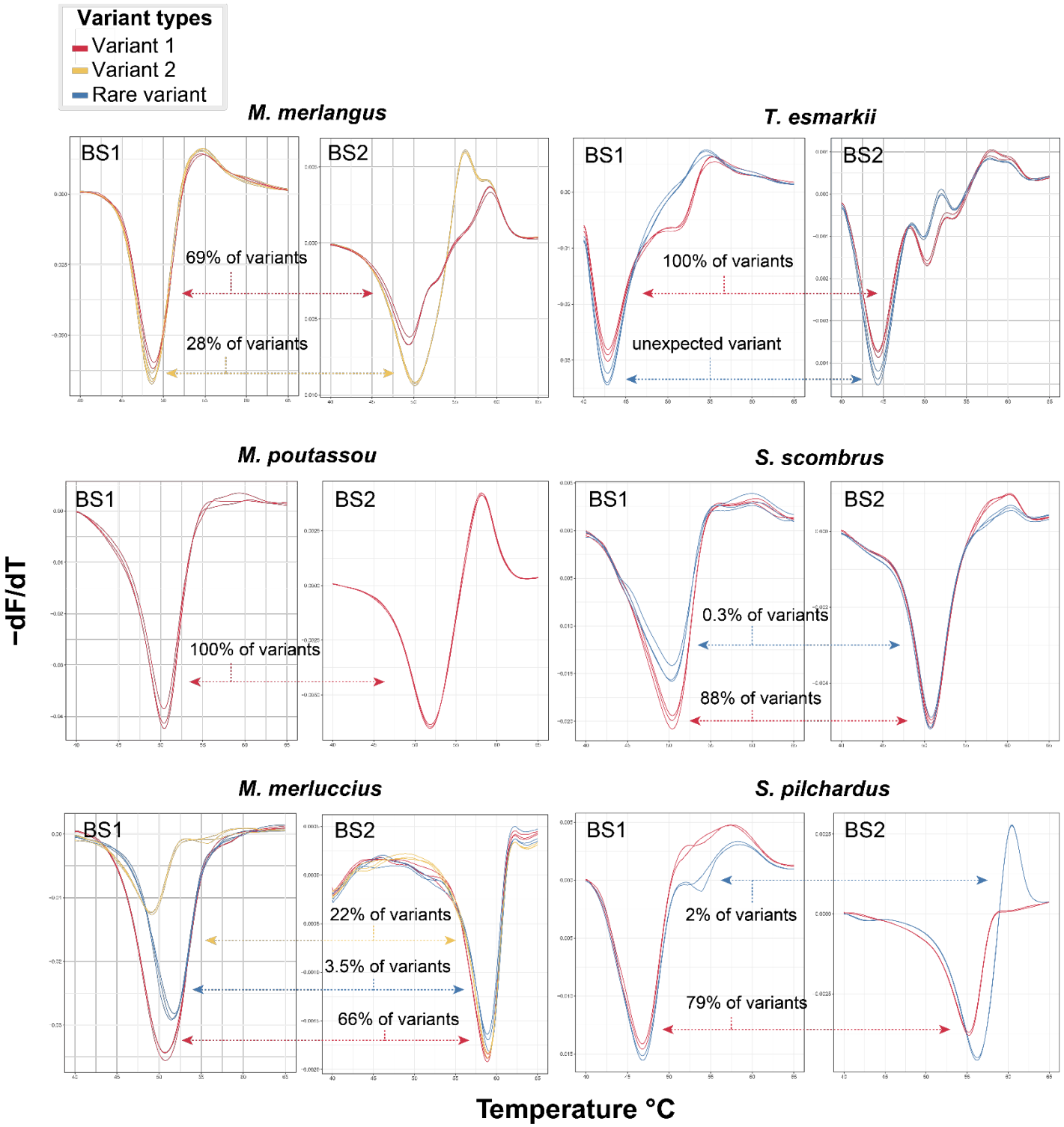


FIGURE 3.4 Illustration of the a) BS1 and b) BS2 species-specific signatures and the intra-specific species variants generated by the voucher samples collected on the RV Cefas Endeavour and processed in a controlled laboratory environment for addition in the online FASTFISH-ID database. For all six species selected, the most common variants were covered.

Identification of field runs against updated reference library

Following the addition of six new species to the online database using a set of new specimens, the Mic run performed on the R/V CEFAS Endeavour were then queried against the updated database to evaluate whether the opportunistically and randomly collected specimens could be unambiguously identified. Five specimens of European pilchard, including one displaying a signature variant were all correctly identified as *S. pilchardus*. On the other hand, the five specimens of Atlantic mackerel (*S. scombrus*) processed on board remained unidentified. Out of the five specimens of Norway pout processed on board, two were recognized by the online algorithm as *T. esmarkii*, while the samples not recognized appeared to display different signatures. Four samples of Whiting displaying two different signatures were correctly identified by the online algorithm as *M. merlangus*, whereas a fifth sample displaying a third signature was not identified. That sample was processed for Sanger Sequencing and contained a single SNP variant on BS2 present only on 1% of Whiting sequences extracted from BOLD (**FIGURE 3.5**). Three out of four samples of Blue whiting (*M. poutassou*) were identified, the fourth sample displayed a different signature variant which was not observed in the voucher samples added to the database. Only one out of two samples of European hake was identified as *M. merluccius*, even though both variants seem to have been added to the database. In total, 15 out of the 26 specimens processed on board from these six different species were recognized by the new online algorithm and none were misassigned or incorrectly matched. Aside from Atlantic mackerel which had no successful identification possibly as a result of contamination or poor amplification of BS1, the online algorithm had on average a 69% success rate at correctly authenticating the remaining 5 species.

Intraspecies variants

To evaluate the extent to which the fluorescent signatures of the specimens collected to populate the updated FASTFISH-ID™ were representative of all possible signature variants, we compared the COI barcode sequence of those specimens for BS1 and BS2 against sequences mined in BOLD. **TABLE 3.1** lists the possible variants identified using BOLD for each of the ten species originally analyzed in a controlled laboratory environment, and the number of specimens we sampled and added to the reference library for each of those possible variants. Based on data mining from BOLD, we expect a minimum of 13 variants, two of which are held by 85% of all individuals for *S. pilchardus*; a minimum of 17 variants, two of which are held by 95% of all individuals for *S. scombrus*; a single variant held by 100% of individuals for *T. esmarkii*; a minimum of 5 variants, two of which are held by 97% of all individuals for *M. merlangus*; a minimum of 16 variants, two of which are held by 79% of all individuals for *M. poutassou*; a minimum of 17 variants, three of which are held by 95% of all individuals and *M. merluccius* (**TABLE 3.1**). Based on the Sanger sequencing results, we were able to establish that among our voucher samples used for building the reference database: Five of the *S. pilchardus* samples covered the most common variant and one of them covered a rare variant; one of the *S. scombrus* sample covered the most common variant and one of the samples covered a rare variant; all three *T. esmarkii* samples covered the most common and only variant expected for that species in terms of sequences but an unexpected fluorescent signature was also observed (**FIGURE 3.4**); two of the *M. merlangus* samples covered the most common variant and one sample covered the second most common variant; all three samples of *M. poutassou* covered the most common variant; and one of the *M. merluccius* samples covered the most common variant, one sample covered the second most common variant, and one sample covered a rare variant (**FIGURE 3.4**).

The sensitivity of the closed-tube barcoding to a single SNP along the barcode segments can be evaluated by visually assessing the variation in the resulting fluorescent signatures. Whiting, for which we expect two dominant variants held by at least 97% of all individuals features a single SNP on BS2 which results in two significantly different

fluorescent signatures for BS2 (**FIGURE 3.5**). Another SNP present in one of the samples processed on the R/V CEFAS Endeavour and displayed in a rare variant also resulted in a different fluorescent signature.

TABLE 3.1 Evaluation of the intra-specific variants for each of the ten species originally selected to completement the FASTFISH-ID™ database based on the COI gene data obtained on BOLD for each species. The number of specimens for which Sanger Sequencing, species identification, and closed-tube barcoding was successfully achieved is displayed for each possible variant. The COI barcode segments from three samples of *D. labrax* and three samples of *C. harengus* either failed to amplify or were not matched to any species, which led to a lack of information regarding the match of the selected specimens to any specific variant.

Scientific name	Number of BOLD specimens with complete BS1 & BS2 sequences	Numer of predicted species-specific signatures	Fraction of signature types	Total percentage of BOLD specimens having each signature type	Number of sampled specimen for each signature type
Merlangius merlangus	93	5	Signature 1 (64/93)	69%	2
			Signature 2 (26/93)	28%	1
			All other signatures	3%	0
Thunnus thynnus	271	8	Signature 1 (256/271)	94%	1
			All other signatures	6%	0
Trisopterus esmarkii	36	1	Signature 1 (36/36)	100%	3
Micromesistius poutassou	96	16	Signature 1 (60/96)	62%	3
			Signature 2 (16/96)	17%	0
			All other signatures	21%	0
Merluccius merluccius	289	17	Signature 1 (190/289)	66%	1
			Signature 2 (63/289)	22%	1
			Signature 3 (21/289)	7%	0
			All other signatures	5%	1
Engraulis encrasicolus	198	40	Signature 1 (65/198)	33%	9
			Signature 2 (48/198)	24%	1
			Signature 3 (22/198)	11%	0
			All other signatures	32%	7
Sardina pilchardus	86	13	Signature 1 (68/86)	79%	5
			Signature 2 (5/86)	6%	0
			Signature 3 (2/86)	2%	1
			All other signatures	13%	0
Scomber scombrus	400	17	Signature 1 (353/400)	88%	1
			Signature 2 (27/400)	7%	0
			All other signatures	5%	1
Dicentrarchus labrax	52	4	Signature 1 - (20/52)	38%	NA
			Signature 2 - (19/52)	37%	NA
			Signature 3 - (10/52)	19%	NA
Clupea harengus	169	8	Signature 1 - (156/169)	92.30%	NA
			Signature 2 - (3/169)	1.80%	NA

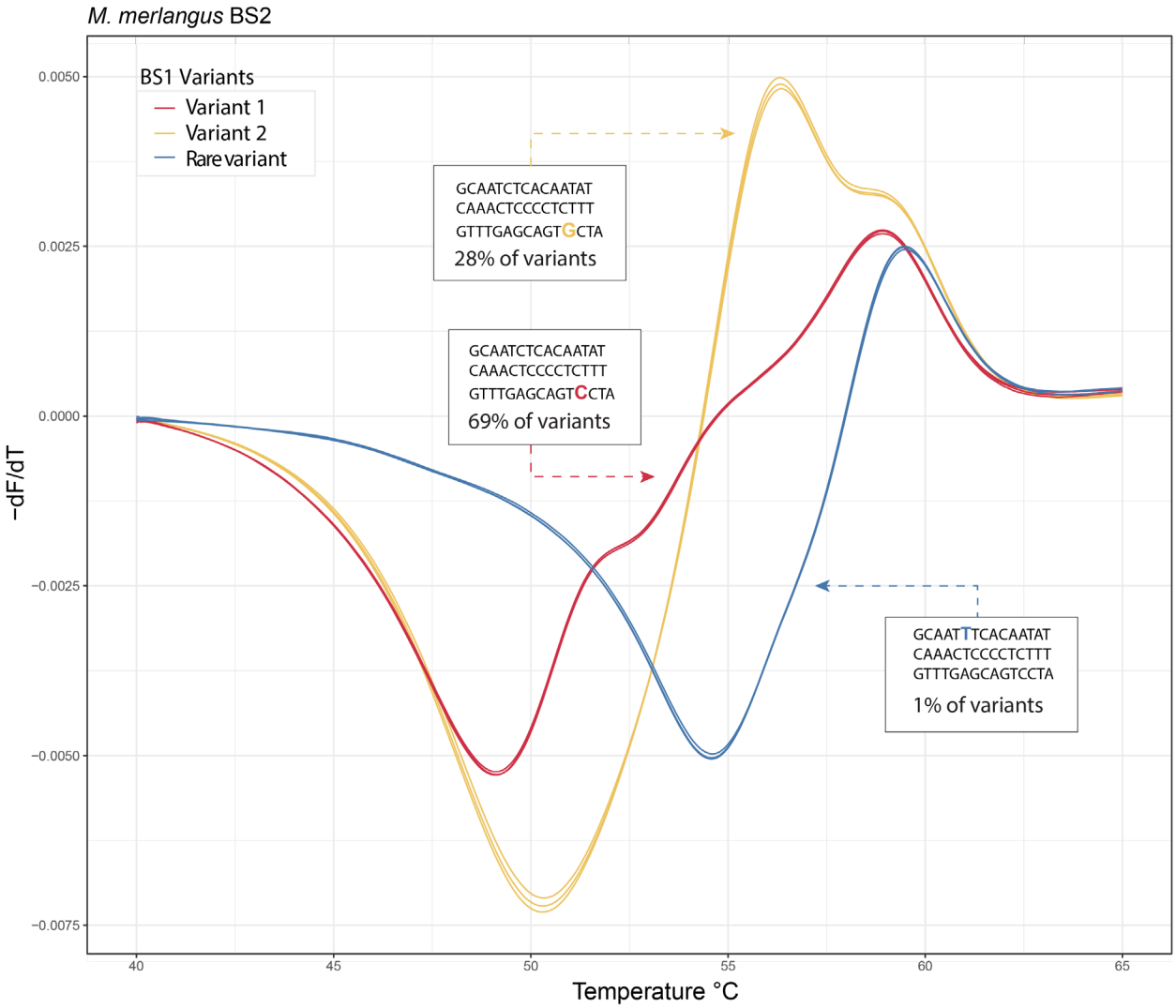


FIGURE 3.5 Illustration of the sensitivity of the closed-tube barcoding technology with *M. merlangus*, allowing to display unique fluorescent signatures given a single SNP in the sequence.

3.4.3 Exploring the use of closed-tube barcoding for intraspecies discrimination

We opted against adding the European anchovy voucher samples to the online database due to the extensive number of variants that exist for that species and to the low percentage of specimens that exhibit the most common signatures (33% of specimens for signature 1 and 24% for signature 2). To have a representative library of all possible European anchovy variants would therefore require a substantial number of samples from

a variety of locations. The Sanger sequencing files from 17 fin samples collected from different individuals aboard the R/V CEFAS Endeavour revealed that within those given samples, one variant dominated, corresponding to a well studied haplotype (haplotype B) which is part of sympatric clades of European anchovy often referred to as haplotype A and haplotype B (Pappalardo et al., 2015). We identified that all of the six COI barcode diagnostic nucleotides for European anchovy haplotype discrimination listed by Pappalardo et al. (2015) were present in the barcode segments targeted by the sets of fluorescent probes designed for the FASTFISH-ID™ method (**TABLE 3.2**). This includes a non-synonymous transversion which results in a different amino acid. The closed-tube barcoding method tested here was therefore able to generate two distinct fluorescent signatures for BS1 and for BS2 (**FIGURE 3.6**) which unambiguously discriminated between these two anchovy haplotypes. Using this set of six diagnostic nucleotides, 13 anchovy samples were identified as belonging to haplotype B, 3 anchovy samples were identified as belonging to haplotype A, and one anchovy displayed a slightly different profile. It should be noted that among the specimens that appeared to belong to haplotype A and B based on screening for diagnostic nucleotides, several specimens (2 specimens of haplotype A and 4 specimens of haplotype B) featured SNPs elsewhere along the barcode segment which led to slightly different fluorescent signatures.

TABLE 3.2 R/V Cefas Endeavour *E. encrasicolus* samples matched against the list of the six nucleotide diagnostics discriminating between haplotype A and haplotype B as described by Pappalardo et al., 2015. Samples highlighted in grey contain one or more SNPs along the barcode segments targeted by the fluorescent probes which may affect the fluorescent signature.

	Nucleotide diagnostics (NDs) present on: BS1																																				BS2			
Reference sequence	**C**	**T**	**T/C**	**T**	**A**	**T/G**	**C**	**A**	**G**	**G**	**A**	**A**	**A**	**C**	**C**	**T**	**T/C**	**G**	**C**	**C**	**C**	**A**	**C**	**G**	**C**	**A/C**	**G**	**G**	**A**	**A**	**C**	**C**	**T**	**C**	**A**	**G**	**T**	**A/G**	**G**	**A**	**T**	**T/C**	**A**
SAMPLE_91_4	-	-	C	-	-	T	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	G	-	-	-	C	-
SAMPLE_110_9	-	-	C	-	-	G	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	G	-	-	-	C	-
SAMPLE_110_7	-	-	C	-	-	G	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	G	-	-	-	C	-
SAMPLE_132_9	-	-	C	-	-	G	-	-	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-	-	G	-	-	-	C	-
Hap. A - Pappalardo et al. 2015	-	-	**C**	-	-	**G**	-	-	-	-	-	-	-	-	-	-	**C**	-	-	-	-	-	-	-	-	**C**	-	-	-	-	-	-	-	-	-	-	**G**	-	-	-	**C**	-
Hap. B - Pappalardo et al. 2015	-	-	**T**	-	-	**T**	-	-	-	-	-	-	-	-	-	-	**T**	-	-	-	-	-	-	-	-	**A**	-	-	-	-	-	-	-	-	-	-	**A**	-	-	-	**T**	-
SAMPLE_91_7	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_132_2	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_43_5	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_43_4	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_110_6	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_43_1	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_91_2	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_31_3	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_132_4	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_91_5	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_110_8	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_43_2	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-
SAMPLE_132_3	-	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-	-	A	-	-	-	T	-

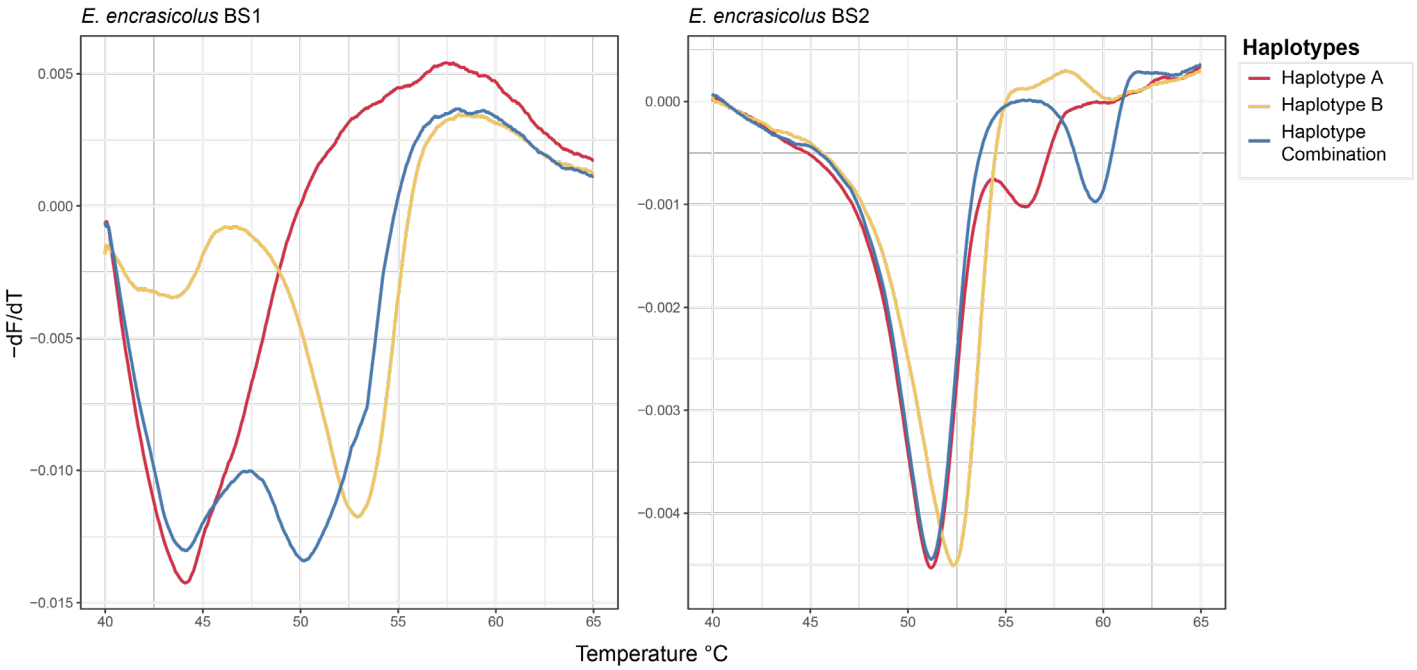


FIGURE 3.6 Anchovy haplotypes identified in Pappalardo et al., 2015 and incidentally profiled on Barcode Segment 1 and Barcode Segment 2 using closed-tube barcoding on specimens collected in the Celtic Sea and English Channel.

3.5 Discussion

We have successfully demonstrated that the closed-tube barcoding method can be used as a rapid and portable *in situ* fish authentication tool when combined with a portable qPCR device such as the Mic. In addition, we showed the relative universality of the FASTFISH-ID™ technology and highlighted how accurate and reliable the method can be by examining the fluorescent signatures and associated barcode segment sequences of randomly selected and opportunistically sampled fish species. Finally, we used the example of the European anchovy to explore and confirm that the FASTFISH-ID™ technology might in some instances not only be well suited to authenticate species but can also perform intra-species haplogroup discrimination.

Successful field applications of closed-tube barcoding

The FASTFISH-ID™ technology has previously been tested in a controlled laboratory setting with carefully selected species as part of a blinded validation study (Naaum et al., 2021) and we here demonstrated that it can easily be transported and performed successfully in the field. The performance of the portable qPCR Mic device did not seem to be affected by unstable conditions such as irregular platform movements due to stormy conditions, which shows that the device itself could easily be transported and moved around, and used in places such as ports, marinas and airports, landing sites and auction markets, as well as moving vehicles like boats, vans, or trucks, provided an electrical outlet or converter is available. This is corroborated by recent studies that have successfully used the Mic for in-field demonstration of coronavirus SARS-CoV-2 identification assays (Paton et al., 2021). In instances where trained staff or specialists are called-in to use a portable device for immediate verification purposes as was the case at the Hong Kong airport with a suspected smuggling event of critically endangered European glass eels (Cardeñosa et al., 2019), a tool that can be readily available and immediately transported to where it is most needed is a non-negligible attribute to such technology. The FASTFISH-ID™ technology, in combination with the Mic, successfully

generated unique species-specific fluorescent signatures on-site for 14 out of the 15 species we randomly sampled, indicating that it could be used as a valuable portable species authentication tool for field applications.

Universality and reliability of closed-tube barcoding for seafood authentication

The added advantage with the closed-tube barcoding technology is its potential universality (Sirianni et al., 2016) which we set out to demonstrate in this study by selecting a random set of commercially important species and evaluating the uniqueness of their fluorescent signatures. Most currently proposed quick and portable DNA-based authentication methods are based on real-time PCR techniques, which have proven to be quite precise and rapid, and which require the development of species-specific primers and/or probes to allow for the rapid detection of species (Fernandes et al., 2020; Silva & Hellberg, 2021). Other emerging portable species identification methods, such as Nanopore DNA sequencing platforms tend to require a higher level of expertise, particularly when it comes to using bioinformatic skills for results interpretation, and if dozens of species are to be multiplexed in a single run, the library preparation can be quite time consuming, expensive, and require a variety of laboratory equipment (Ho et al., 2020). As discussed by Naaum et al., (2021), aside from its ease of use, the appeal of the FASTFISH-ID™ technology is its universality, and in their controlled trials on 75 specimens from 18 commercial fish species, the authors reported a 98% success rate in correct authentication, and further revealed no instances of incorrect identification. Given the implications of such claims and the value this could represent for the fisheries industry and control entities, it was remarkable that, among randomly selected species that demonstrated successful amplification of BS1 and BS2 in the field, the present study achieved an average authentication success rate of 69% with no instances of misassignment. This success rate varies greatly between species, with some such as European pilchard displaying a success rate of 100% while others, such as Norway pout, displaying a success rate of 40% due to variants. These results highlight that the method is universal, but that successful authentication is dependent upon the integration of all

species variants in the database, which was a limitation of this study. Variants could easily be synthetically manufactured and included in the reference database without the need to sample voucher specimen.

In most instances where such a technology would be used other than for research purpose only, such as law enforcement or prosecution, reliability is a non-negotiable key component (Martinsohn et al., 2019). Despite relatively low authentication success rate for some species due to the study limitations leading to an incomplete database, we did not note a single instance of misassignment, which is a crucial component of such a universal method. The reliability of the closed-tube technology and the FASTFISH-ID™ assay seems therefore to be quite high. On the contrary, as suggested by Rice et al. (2012) we noted that the method is highly sensitive even to changes in a single nucleotide along the targeted barcode segments, which could explain some of the low success rates observed, as any sequence variants need to be accounted for and added in the species reference library.

Intra-specific variation and multiple intra-species fluorescent signatures

The high sensitivity of this technology compared to other DNA-based methods such as Sanger sequencing can both be an advantage and an inconvenience. It is an advantage as it can tell apart species with minimal genetic differences within the target barcodes, but it comes as a disadvantage in instances of intra-species variation or heteroplasmy. For example, despite similar sequences as identified by Sanger sequencing, the Norway pout results highlighted here displayed two slightly different fluorescent signatures (**FIGURE 3.4**). This could be an instance of heteroplasmy that has not been detected through Sanger sequencing but that was picked up by the fluorescent probes leading to slight denaturation temperature profiles.

The barcode segments targeted by the fluorescent probes were originally chosen as they present reduced within-species variation and maximize between-species variation (Naaum et al., 2021). It is however difficult to find barcode segments where variants will

not exist within a species, and therefore, these variants need to be accounted for when using the closed-tube barcoding technology. An advantage of the FASTFISH-ID™ assay is that it uses barcode segments situated in the COI barcode, allowing for easy scrutinizing of COI barcode sequences through universal databases such as BOLD and NCBI. In reviewing the BOLD database for each of the species examined here, we noted that some species such as Atlantic mackerel or blue whiting display up to 16 different variants which would themselves theoretically result in 16 different signatures for either one of these species. Though a species can display intra-specific variant heterogeneity, it is important to note that some variants are much more common than others. Based on data for 390 specimens of Atlantic mackerel added on BOLD for example, a single variant is accounting for about 88% of the total number of variants observed. Based on these ratios, we therefore evaluated that most of the species we selected contained the most common variants expected for any given species. As demonstrated with the samples of blue whiting tested in this study however, for which the two most common variants were added to the database, a lack of identification of one specimen was here linked to not having added one of the rare variants to the database. This illustrates the need to incorporate more individuals or more variants in the database, including rare ones, to avoid the risk of running into unidentifiable specimens, which can easily be done by creating synthetic oligos of all possible variants.

Some specimens failed to be identified either as a result of poor amplification or possible variants. None of the Atlantic mackerel processed onboard the R/V CEFAS Endeavour were identified and given that one of the voucher specimens used for building the library displayed the most common variant and that both specimens came from the same population as the specimens screened for authentication, it is unlikely that the failed identification was due to variants. Instead, it seems like the BS1 from the specimens processed on board was not adequately amplified leading to a lack of identification. In general, most of the specimens that failed to be identified simply displayed a different variant, one specimen of European hake however was not identified despite displaying a variant that was added in the database. This could simply indicate

that the algorithm used to assign specimens to their correct species might benefit from having more than one specimen per variant, which was not the case here.

Highly variable sequences and haplotype discrimination

European anchovy generated a diverse set of signatures suggesting that either the target barcode sites are quite variable, or individuals might be prone to heteroplasmy, or samples had been contaminated. Upon examining Sanger sequencing results, completing further Mic runs with muscle tissue instead of fin clips, and obtaining more consistent signatures, it seemed like much of the variation observed was due to sequence variants, though some may have also been due to intraspecific contamination of the fins. Exogenous (non-target) DNA can contaminate external tissue of broadcast spawners and schooling fish such as fins, leading to erroneous results (Petrou et al., 2019). Results from Sanger sequencing clearly highlighted that a difference in curve amplitude should not be considered a variant but that a shift in the peak or the trench observed on the fluorescent signatures of different individuals should be interpreted as resulting from one or several SNPs.

The COI barcode segment of the European anchovy is particularly variable, and the barcode segments targeted by the FASTFISH-ID™ probes contain a minimum of 40 different variants with the dominant variants accounting only for a small percentage of the total variants. These observations render this species an unlikely candidate for this technology, unless different barcode segments with less variation were targeted by the probes thereby defying the purpose of a “universal” method. That being said, European anchovy’s mitochondrial DNA’s variability has been extensively studied (Magoulas et al., 2006; Pappalardo et al., 2015; Viñas et al., 2014; Zarraonaindia et al., 2012) revealing the existence of populations that could be identified simply with a few SNPs from the mtDNA control region (Viñas et al., 2014; Zarraonaindia et al., 2012). Two clades of European anchovy often referred to as haplogroups A and B have been identified (Silva et al., 2014) and it is possible to discriminate between these haplogroups with a few diagnostic SNPs

present in the COI barcode region of the mtDNA (Pappalardo et al., 2015). These SNPs unexpectedly happen to all be part of the two barcode segments selected by the probes developed for the FASTFISH-ID™ assay, generating unique fluorescent signatures that can be attributed to either haplogroup A or haplogroup B. Studies have demonstrated that these clades seem to be associated with latitudinal thermal conditions (Silva et al., 2014) with haplogroup B being more prevalent in higher latitudes in the Northern hemisphere. Among the samples analyzed in this study, haplogroup B was indeed the most prevalent clade.

The survey from which these samples were collected is organised yearly by the UK's Department for Environment, Food and Rural Affairs, and aims at evaluating the stock conditions of pelagic species in the Celtic Sea, Irish Sea, and English Channel with a large focus on European Anchovy. We here unexpectedly demonstrated that the FASTFISH-ID™ method could successfully provide an additional layer of information whereby visually identified specimens of European Anchovy could then be divided into clades. With a method like this one, which combines closed-tube barcoding with a portable device capable of processing up to 48 samples in two hours, the survey could process several dozens of samples a day and easily map the ratio of these haplogroups, evaluating throughout the years how certain environmental phenomenon such as climate change might affect the presence of these clades. This may also apply to other species, in other climatically dynamic regions.

3.6 Conclusions

We here not only demonstrate the rapidity, practicality, and portability of the closed-tube barcoding FASTFISH-ID™ method in the field but also its relative universality and potential use for haplotype discrimination opening a wide panel of new applications for this method. Molecular tools are particularly handy in instances where taxonomic identification is difficult or impossible. At the moment however, robust tools that prove easy, quick and portable also require the species-specific customization of assays which is an impediment to their use and application. Given how problematic seafood fraud can be for the sustainable management of fish stocks, commercial and enforcement bodies as well as non-governmental organisations have long been demanding seafood authentication tools that are portable, easy to use, and rapid. The fight against illegal trade of wildlife products in general can largely benefit from molecular tools provided they are reliable and cater to the needs of interested parties. Promising tools such as the FASTFISH-ID™ technology must therefore place themselves on the forefront of possible forensic methods used for enforcement and prosecution while also adapting to the budget, time, and expertise constraints of stakeholders and enforcement entities.

Chapter 4

A future for seafood point-of-origin testing using DNA and stable isotope signatures

This chapter was published in *Reviews in Fish Biology & Fisheries*

Authors: **Cusa, Marine***; St John Glew, Katie*; Trueman, Clive; Mariani, Stefano; Buckley, Leah; Neat, Francis; Longo, Catherine

* These authors contributed equally to this work as first authors: Marine Cusa was in charge of the genetics aspect of the study, and Katie St. John Glew was in charge of the stable isotopes aspect of the study.

Important note on this chapter:

The manuscript generated as part of this chapter was developed as a first-co-author paper. Some text was written, edited, and remains as part of a collaboration between the two first authors. In the methods and the results sections, I have only included the work for which I was responsible. Reference to the manuscript can be found in the Supplementary Material.

4.1 Abstract

Demand for seafood products is increasing worldwide, contributing to ever more complex supply chains and posing challenges to trace their origin and guarantee legal, well-managed, sustainable sources from confirmed locations. While DNA-based methods have proven to be reliable in verifying seafood authenticity at the species level, the verification of geographic origin remains inherently more complex. Both genetic and stable isotope analyses have been employed for determining point-of-origin with varying degrees of success highlighting that their application can be effective, when the right tool is selected for a given application. Developing an *a priori* prediction of their discrimination power for different applications can help avoid the financial cost of developing inappropriate reference datasets. Here, we reviewed the application of both techniques to seafood point-of-origin for 63 commercial finfish species certified by the Marine Stewardship Council, and showed that, even for those species where baseline data exist, real applications are scarce. To fill these gaps, we synthesised current knowledge on biological and biogeochemical mechanisms that underpin spatial variations in genetic and isotopic signatures. Here we describe which species' biological and distribution traits are most helpful in predicting effectiveness of each tool. Building on this, we applied a mechanistic approach to predicting the potential for successful validation of origin to three case study fisheries, using combined genetic and isotopic methodologies to distinguish individuals from certified versus non-certified regions. Beyond ecolabelling applications, the framework we describe could be reproduced by governments and industries to select the most cost-effective techniques.

Keywords

Authentication · Chain of custody · Geographical origin · Mislabelling · Traceability · Validation

4.2 Introduction

Increasing demand for seafood from a growing global population generates concerns over sustainable exploitation and mitigation of environmental impacts of fishing activities (FAO, 2020). Excluding products coming from unsustainable sources from the market is a way to support well-managed fisheries and to help remove incentive for poor or illegal practices. This is the intent, for example, of legally mandated checks at landing sites by signatory States to the FAO's Port State Measures Agreement (OECD, 2018). This, however, remains the first stage in an often lengthy seafood supply chain. Once a fish is landed and starts its journey through the supply chain, it becomes increasingly difficult to track whether it originated from legal and sustainably managed fisheries as it proceeds through the far-reaching and complex global seafood trade networks (FAO, 2020; Leal et al., 2015; Yasuda & Bowen, 2006). To guarantee the integrity of the chain of custody, each step throughout the supply chain must be documented and fully traceable. Further checks may be needed at import border crossing, processing plants, and key points in the supply chain. Yet, traceability tools, which can be defined as the methods used to follow a product along the supply chain such as landing declarations, catch certificates, supplier self-reporting, volume reconciliation, etc., are vulnerable to manipulation. Illegally caught fish resulted in a worldwide loss of US$ 10-23.5 billion in 2009 as estimated by Agnew et al., (2009) and of US$ 9-17 billion in 2020 according to (Sumaila et al., 2020). Authorities checking for compliance, as well as seafood businesses interested in protecting their brand or passing denomination of origin or ecolabeling audits, require diagnostic tests to confirm provenance (i.e. the geographical point of origin) documentation.

Forensic point of origin testing is a key tool for verifying traceability information (Ogden & Linacre, 2015) and it is becoming standard practice within many industries and for products such as meat, dairy, wine, and honey (Donarski & Heinrich, 2015; *FERA*, 2020; Kelly, S., Heaton & Hoogewerff, 2005; Morin & Lees, 2018; RedTractor, 2013). DNA profiling, stable isotope analysis, fatty acids, and elemental profiling, have all been tested to varying degrees with shellfish and finfish populations of farmed and wild caught seafood (Gopi et al., 2019). Due to the ease with which these tracers can be used in

terrestrial systems, genetic and biogeochemical markers are particularly well-established as tools to verify breed and region of origin of food products in terrestrial food chains (Heaton et al., 2005, 2014; Kelly, S., Heaton & Hoogewerff, 2005) but are far less frequently used in marine systems and, we argue here, deserve more attention.

Genetic and biogeochemical tracers are mechanistically and analytically different and can vary independently. This presents both difficulties and opportunities, potentially increasing the value of combined approaches. Stable isotope, fatty acid and elemental tracers for spatial origin are based on the transmission of the tracer signal from the underlying natural environment into the organism (Ramos & González-Solís, 2012). These techniques therefore directly link an organism to a physical location at a point in time, but their efficacy is dictated by underlying spatial gradients in the tracer signals. Genetic tracers are instead grounded in the fundamental processes of inheritance and are therefore based on the dynamic interplay of isolation/exchange of allelic variants existing within and among reproductive groups (J. Freeland et al., 2005; Wright, 1931). Genetic tracers therefore reflect population spatial histories rather than recent individual movements or location at the point of capture.

Forensic provenance testing will generally require the development of a reference database of known origin samples to be collected across the regions of interest for any given species (Kelly, S., Heaton & Hoogewerff, 2005) and this is often logistically and financially challenging. Provenance tests are considerably more difficult to validate across large and often poorly sampled marine environments compared to agriculture, farming, or even aquaculture, where species are often spatially constrained. In addition, since marine environments have fewer physical boundaries, individuals and populations of fish species move and migrate to varying extents and mix with other populations. The accuracy and precision associated with natural tracers used as markers of geographic origin will depend on the nature of spatial variance in the tracers in question, and on the quality of the reference dataset (**FIGURE 4.1**).

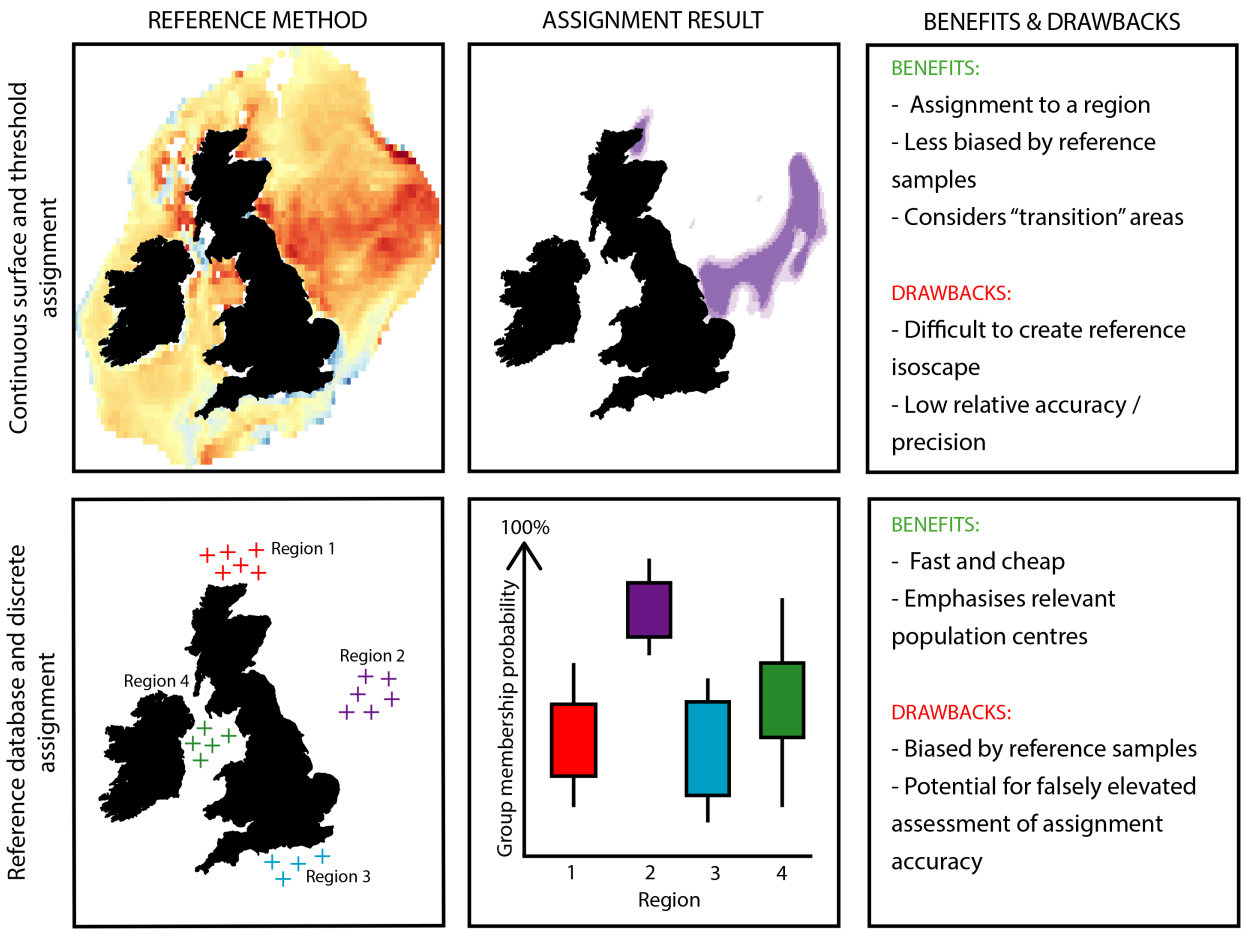


FIGURE 4.1 Conceptual schematic depicting the differences between discrete and continuous assignment methods and associated reference materials required. Continuous assignment approaches (top row) rely on an even coverage of data over the entire study area, interpolated through a spatial model across a grid of varying resolution dependant on data availability (top left panel), and provide the highest likelihood of origin area (coloured area in the top right panel), compared to all other possible locations (ie other grid cells). Discrete assignment approaches (bottom row) require a discrete set of samples from each of the areas of interest (sample locations identified by crosses for each of four regions, lower left panel) and provide the probability of sample assignment to each of these predetermined areas (lower right panel). Benefits and drawbacks in terms of practical applicability and detection power of each approach are highlighted.

Developing an a priori prediction of the power for any given spatial marker technique to discriminate among specific fishery populations can help avoid the financial cost of developing inappropriate reference datasets. At the same time, this exercise can potentially open spatial verification methods to a wider range of users and contexts. Here we present case studies for which two very different types of currently available tools, genetic and stable isotope tracers, can be used to establish seafood provenance at spatial scales and accuracy measures relevant to real world applications. Genetic tools

are already commonly used in seafood forensics for species identification, though their use for provenance testing is less advanced. Stable isotopes are a good model biochemical tracer due to the extensive research and use within the terrestrial food traceability sector (Chesson et al., 2008, 2010; Kelly, S., Heaton & Hoogewerff, 2005) and the successful use in discriminating origin in wild caught seafood products at both local and broad geographical scales (Carrera & Gallardo, 2017; Gong et al., 2018; Kim et al., 2015). In addition, the growing availability of varying spatial isotopic marine models resolution (Trueman & St John Glew, 2019) provides an opportunity to predict where further traceability case studies are likely to succeed. We focus on commercial fisheries certified against the Marine Stewardship Council (MSC) Fishery Standard, expecting them to have good availability of biological and traceability research. In order to exhibit the ecolabel on consumer-facing products, the MSC program requires chain of custody certification for each step of the supply chain and must assure that products were harvested in the location and by the fishers covered by the certificate. The MSC example therefore represents a useful proof of concept as it combines a claim of origin from a sustainable stock and defines the unit of certification to a particular region and group of harvesters.

Through a meta-analysis we review the extent to which species biological traits can predict genetic population structure. We develop a workflow to estimate the likely efficacy of either stable isotope or genetic tools when applied to a defined spatial verification problem to enable us to investigate the capabilities of these techniques using stocks certified in specific locations by the MSC. Beyond the ecolabeling examples presented, the provenance testing protocol and methods described here can be applied to a range of provenance-testing questions for any species of both fish and shellfish within a marine environment.

Box 1

Stable isotope and genetic assignment methods

Stable isotope assignment:

Isotope ratios vary spatially across terrestrial, aquatic and marine landscapes due to varying environmental conditions exerting effects on isotopic abundance (Bowen, 2010), due to differences in the hydrological cycle, fluid dynamics, nutrient cycling and biological processes. To utilise this spatial isotopic variation, or determine values in unmeasured regions, isotope maps (isoscapes) are produced. In marine environments, organic isotopic composition varies across space in phytoplankton at the base of the food web due to differences in rate of photosynthesis, and the nutrients available within the water column (McMahon et al., 2013; Ramos & González-Solís, 2012). These spatial isotopic variations are then transferred up the food web, enabling isotopic ratio measured in the tissue of a fish to act as a natural tag (Ramos & González-Solís, 2012), indicating the individual's foraging location before capture.

Genetic assignment:

Genetic methods can readily discriminate between most species owing to the long-standing reproductive isolation between evolving lineages, resulting in measurable DNA sequence divergence (Hebert, Cywinska, et al., 2003). Complications arise when investigating the geographical origin of a specimen, as this requires the given species to be composed of somewhat reproductively isolated populations, and that sufficient, detectable genetic variance exists among these groups. A population can be defined as a group of individuals from a given species living in a set geographical area, interbreeding, and displaying some degree of reproductive isolation from other populations (Freeland & Petersen, 2011; Waples & Gaggiotti, 2006). In nature however, populations are not always fully reproductively and/or geographically isolated. The population concept can be visualized as functioning on a continuum with various degrees of connectivity, from total panmixia where individuals of reproductive age are effectively mating randomly with each

other, to complete isolation where reproduction between populations is impossible (Waples & Gaggiotti, 2006).

Box 2

Definition of verification and assignment

The two main uses of forensic provenance testing methods are for 'verification' or 'assignment' purposes (Nielsen, 2016). Here we define 'verification' as the use of spatially varying natural markers to test the likelihood that a specific geographic claim is true, e.g., checking if the seafood actually came from the region stated on the product's label. 'Assignment' refers to the use of natural markers to infer location or origin, e.g., finding out where an unlabelled fish has come from. Assignment may be based on discrete approaches, where the sample of interest is matched against a set of reference samples chosen to characterise previously defined possible source areas, or continuous approaches where the reference data are transformed into a continuous probability surface or map, using a model to fill in information from areas where reference samples were not present (**FIGURE 4.1**). The requirements for reference datasets increase from verification (a selection of samples from the predicted region of origin), to discrete (a broad selection of samples from the areas to be distinguished between), and continuous assignment designs (an evenly spatially gridded set of samples across the entire region of interest). Verification is overwhelmingly the most common design for natural tracer studies in food forensics and traceability applications.

4.3 Methods

Secondary source data collection

In total, we focussed on 74 marine finfish, 11 of which were in assessment and 63 of which were certified against the MSC Fishery Standard and covered under 133 different certificates (Link to data on Dryad[1]) as of November 2018. An MSC certified fishery is defined here as the group of vessels operating under an MSC certificate in a particular area with a particular gear and targeting a particular species (MSC, 2020). For example, 17 different fisheries are certified to fish Atlantic cod (*Gadus morhua*) in the NE Atlantic. This list excludes salmon due to their anadromous life cycle which makes them akin to freshwater species in terms of reproductive isolation.

We first conducted a literature review to evaluate some of the distinct life-history traits that have previously been identified as important in influencing the outcome and interpretation of both genetic studies and stable isotope analysis (**TABLE 4.1**) (Link to data on Dryad). We then collected information on these traits for each species included in the study through another set of extensive literature review, and later gathered data on species-wide population genetic structure, in the sub-set of species for which this information was available, using a widely employed indicator of genetic dissimilarity (the F_{ST} value (Wright, 1965)). We only collected global F_{ST} or average F_{ST} values and, when possible, collected several F_{ST} values for any given species. Given that F_{ST} values for a single species may differ depending on the geographical range covered by the study, we estimated this geographical coverage for each study, and labelled it as "Entire", "Substantial", or "Regional" coverage.

Analysis of the secondary sourced genetic data

Using attributes in **TABLE 4.1**, we populated a database including, for each listed finfish species, the behaviour and life-history traits relevant to population genetics (e.g., migration mode, larval dispersal potential, etc., see **TABLE 4.1**), and the corresponding

[1] https://datadryad.org/stash/share/ZQ8_3oQ17Kt8mWXJ3-gGgq8pczuRoX9a6I4DzmkuRD4

genetic structure information, when available. Not all these life history categories have well defined quantitative thresholds, and we therefore used the following definitions to evaluate attributes for each species: Migration (None = not migratory, Limited = displays some level of migration but too low to be considered migratory per se, Migratory = clear, well-studied migration patterns for foraging or reproduction, Highly Migratory = large intercontinental migrations); Habitat (Pelagic, Benthopelagic, Benthic – as observed and noted in the literature); Distribution (Only continental margin = only one continuous continental margin, beyond continental margin, worldwide – as observed and noted in the literature); Depth Zone (Aphotic, Euphotic, Disphotic – based on the average depth range of adults as observed and noted in the literature); Larval Dispersal Potential (A comprehensive and qualitative evaluation of larval dispersal potential which includes larval pelagic duration, larval type, larvae buoyancy, diel vertical migration behaviour, larval homing behaviour and swimming abilities, as well as observed or modelled advection patterns, based on literature evidence). We used a combination of FishBase, primary literature, and MSC public certification reports to obtain information on the life-history attributes of the species. We then applied a multivariate ordination technique to visualise these data and allow graphical representation in a two-dimensional space. Due to the categorical nature of the life-history variables and to their mutually exclusive nominal levels (Eg. Distribution [Low, Medium, High], Migration level [Low, Medium, High], larval dispersal potential [Low, High]), we used a Multiple Correspondence Analysis (MCA) which is well suited to multilevel categorical, rather than continuous, variables (Abdi & Valentin, 2007). For this analysis, we removed species for which life history trait information was missing to avoid missing values bias that could affect species correlations.

To evaluate the effects of life-history traits on population genetic structure, we used genetic structure estimates, i.e., F_{ST} values, as the response variable in our statistical analyses. We only retained marine fish species for which Wright's F_{ST} index values were available with the use of microsatellites or Single Nucleotide Polymorphisms (SNPs). Due to their viviparous reproduction method and to the lack of information available on their

early life history, species belonging to the genus *Sebastes* were not included in this analysis. We first used an Analysis of Variance (ANOVA) to test whether the level of genetic structure was significantly different between teleost families. We log-transformed F_{ST} values to meet parametric test assumptions and approach normality, we then computed the homogeneity of variance across groups using Levene's test and the residuals normality using Shapiro-Wilk test. In order to control for some of the elements inherent to such meta-analysis and that may bias the results, we then used a linear model to test for the effect of marker types and geographical coverage on F_{ST} values before evaluating which attributes might best predict population structure. Marker types did not significantly affect F_{ST} values, but geographical coverage did, with 'substantial coverage' displaying significantly different F_{ST} values than 'entire coverage' (**FIGURE 4.2**). Based on these results and since geographical coverage is not a biological factor but rather an artefact of study methodology and quality, we conducted the rest of the analysis separately for studies that covered the entire distribution of the species and for studies that covered a substantial amount of their distribution. We removed studies that encompassed a regional coverage of a species distribution as we only had data for four of them and as they are less likely to be representative of overall species population structure. Finally, we conducted several linear models to identify whether trait combinations affected population structure on the log transformed F_{ST} values, which allowed for better approximation of normality of F_{ST} values. We then attempted to correct for the selection of over-parametrised models by accounting for the adjusted coefficient of determination and for the Akaike Information Criterion (**TABLE 4.2**). We ran linear models separately for studies covering different geographical coverage, as geographical coverage significantly affected F_{ST} values whereas marker types did not (**FIGURE 4.2**).

We performed the MCA analysis using R version 3.6.0 (R Core Development Team, 2018) via the FactoMineR (Lê et al., 2008) and factoextra packages (Kassambara & Mundt, 2017). We also performed the Linear Models and Generalized Linear Models using R version 3.6.0 (R Core Development Team, 2018). We generated the figures using ggplot2 (Wickham, 2016) and ggpubr (Kassambara, 2020).

TABLE 4.1 Ecological attributes relevant to population genetic structure investigated for each MSC certified marine fish species.

Ecological Attributes	***Categories***			
Migration	None	Limited	Migratory	Highly migratory
Habitat	Pelagic	Benthopelagic	Benthic	
Distribution	Only CM*	Beyond CM*	Worldwide	
Depth zone	Aphotic	Euphotic	Disphotic	
Larval dispersal potential	Low	High		

** CM = Continental Margin*

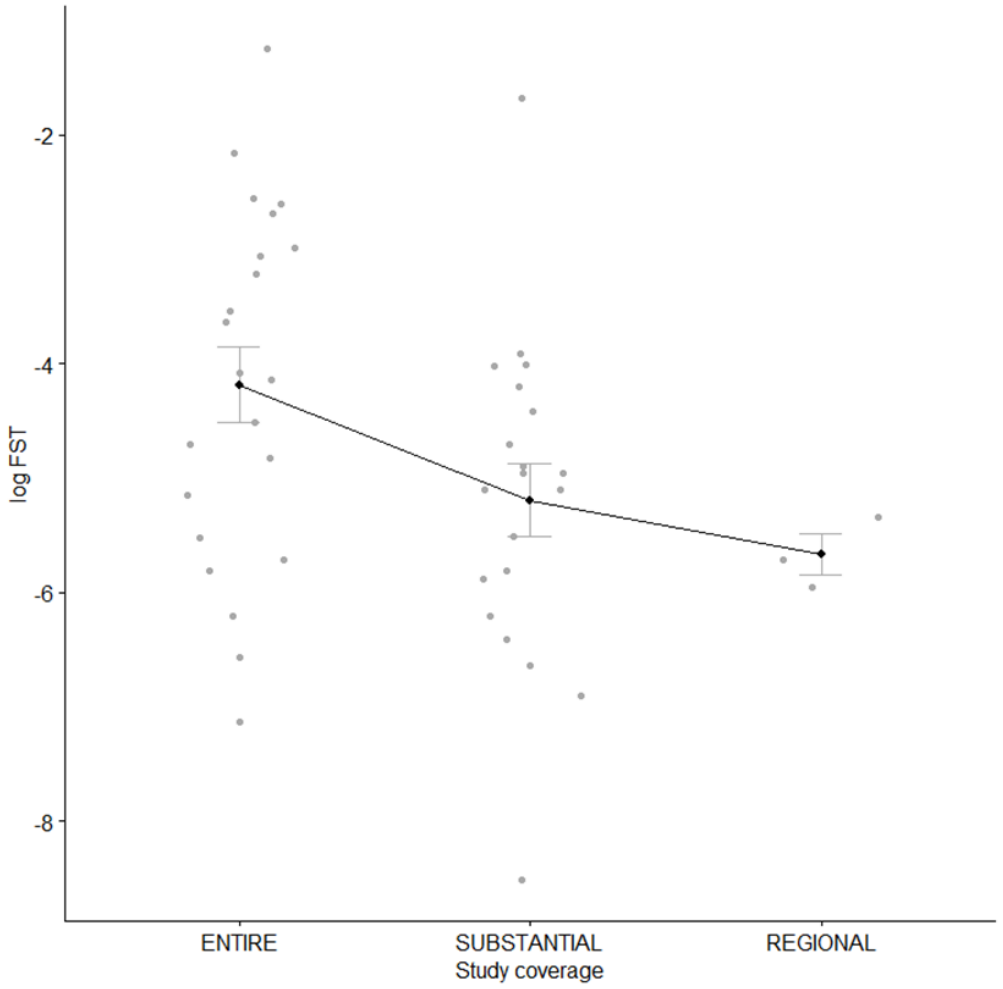


FIGURE 4.2 The effect of study geographical coverage scale on log F_{ST} values. Studies were assigned three types of coverage; "Entire" if they covered all of the species geographical distribution, "Substantial" if they

covered a majority of the species geographical distribution but not all of it, and "Regional" if they covered a restricted area of the species' overall geographical distribution. ($F = 3.138$, $df = 2,42$, $p = 0.0313$, with Levene's test indicating homogeneity of variance ($F = 2.17$, $p = 0.13$) and Shapiro-Wilk test indicating normally distributed residuals ($W = 0.99$, $p = 0.97$).

4.4 **Results**

Genetic structure

The MCA grouped organisms according to the life-history features they share (**FIGURE 4.3**) (**Table S4.1**). The first two dimensions account for 31.13% of the variation observed. The MCA factor map illustrates an off-centred cluster on dimension 1 composed of several tuna species and the swordfish along with the levels indicating, as expected, highly migratory species with a worldwide distribution. As illustrated by the habitat colour pattern, most pelagic species are found on the positive side of the first dimension whereas most benthic and benthopelagic species are found on the negative side of the first dimension. These species are separated along the second dimension based on their distribution (beyond continental margin, continental margin only), and migration patterns. Thus, species like Atlantic cod (COD), saithe (POK), and Greenland halibut (GHL) are clustered on the positive side of dimension 2 describing species that are benthopelagic, migratory, and distributed beyond the continental margin; whereas several species of sole (LEM, RFE, SOL) and rockfish (RFC, SBC, SGO), are clustered on the negative side of dimension 2 and 1, describing species that display limited migration, preference for benthic habitats, and/or that only occur on the continental margin.

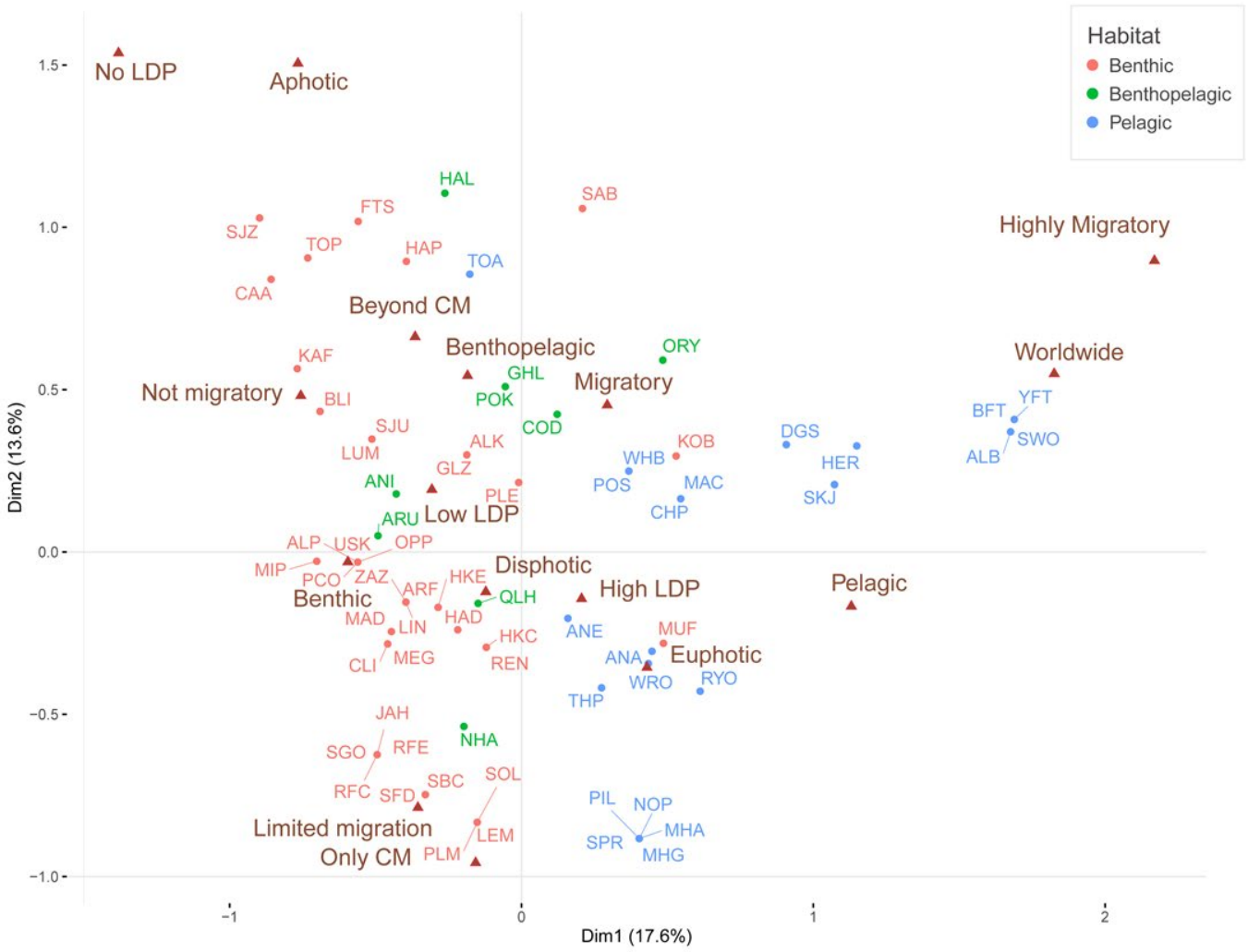


FIGURE 4.3 Projection of the first two dimensions of a Multiple Correspondence Analysis illustrating clusters of fish species based on ecological traits. The three letter codes correspond to 70 different fish species which are listed in **Table S4.1**.

Population genetic structure information was collected for 32 marine fish species belonging to 15 different families. The families that contained the highest number of species were the Clupeidae with six species, the Gadidae with six species, and the Scombridae with four species. An ANOVA indicated that family is a significant factor affecting F_{ST} values (**FIGURE 4.4**). Clupeidae displayed a clustered pattern with relatively and consistently low F_{ST} values (ranging from 0.002 to 0.018), this was also true for Lotidae

(ranging from 0.0014 to 0.0061), and aside from one species (*Gadus morhua*, with an overall F_{ST} value of 0.051), all other Gadidae (ranging from 0.015 for Pacific cod, *Gadus macrocephalus* to 0.003 for saithe, *Pollachius virens*). On the other hand, Scombridae displayed a wider range of F_{ST} values (ranging from 0.003 to 0.1).

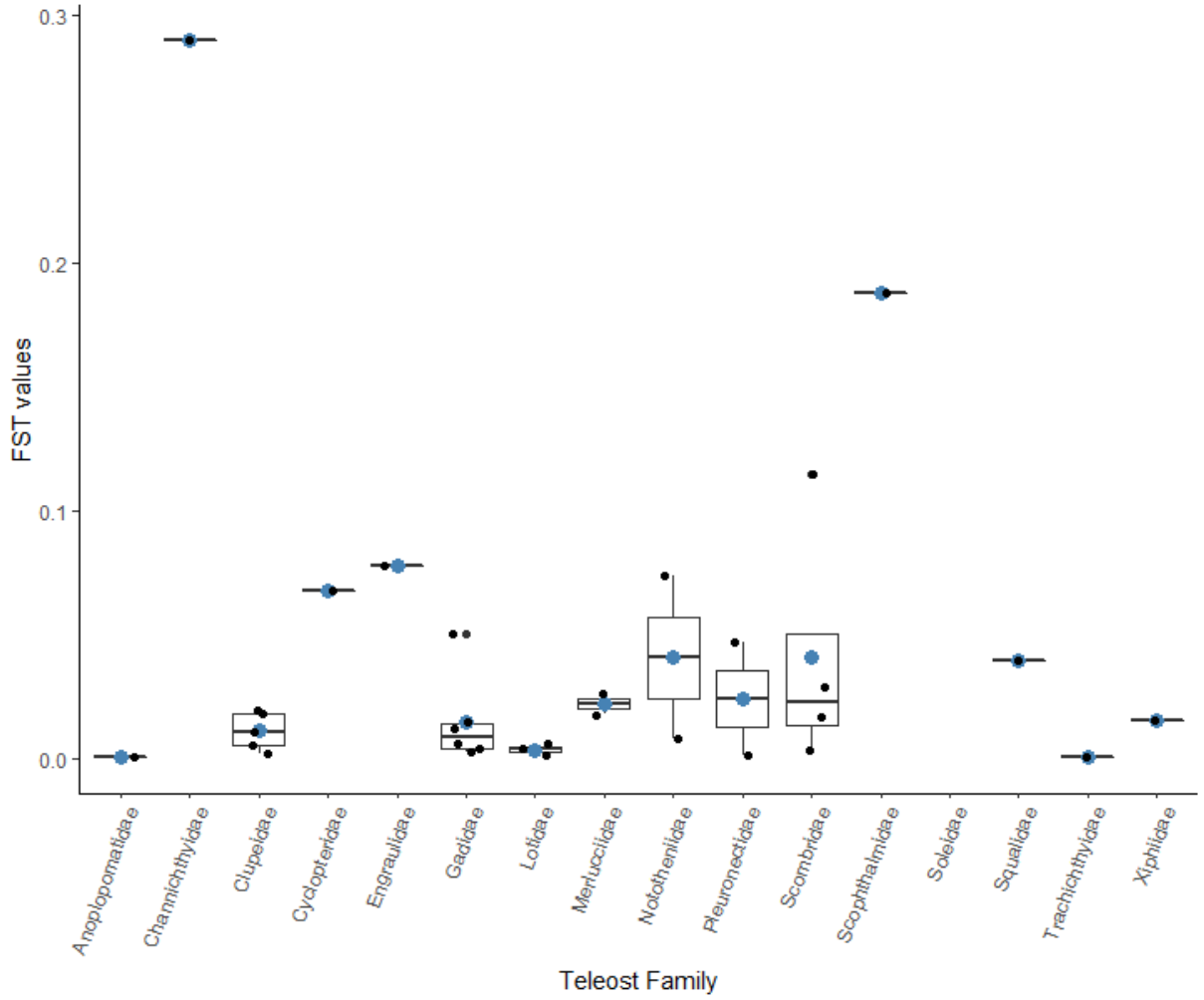


FIGURE 4.4 Non-transformed F_{ST} values for 32 fish species from 15 marine fish families highlighting how some species from a given family may cluster into low-F_{ST} value groups, whereas other families may contain species that display a wide range of F_{ST} values. F = 2.362, df = 14, 17, p = 0.0473 ,with Levene's test indicating homogeneity of variance (F = 1.45, p = 0.23) and Shapiro-Wilk test indicating normally distributed residuals (W = 0.96, p = 0.21).

When focusing on the secondary sourced data from studies that covered a substantial amount of the species distribution, larval dispersal potential significantly affected population genetic structure values in all models. A linear model with larval dispersal potential as a single factor accounting for 23% of the model variance suggests a positive relationship between low larval dispersal and increasing log F_{ST} values. This phenomenon was not observed for the secondary sourced data from studies that covered the entire distribution of the species. The model that best fitted the data for both substantial coverage and entire coverage studies included larval dispersal potential, habitat, migration, and depth zone as explanatory variables but did not include distribution (**TABLE 4.2**). Those best fit models evaluated based on maximum adjusted r^2 and on lowest AIC values indicated that species with low larval dispersal and no-migratory behaviour are likely to exhibit highest log F_{ST} values, corresponding to higher population genetic structure. Pelagic species also had a positive relationship with log F_{ST} values. Depth zone presented some conflictual results suggesting a positive relationship between log F_{ST} values and surface-dwelling habits (disphotic and euphotic zones), for the substantial coverage studies, and negative relationship between log F_{ST} values and a euphotic zone dwelling habitat for entire coverage studies. This model was significant and explained a high proportion of the observed variance for both the substantial coverage and the entire coverage datasets (75% and 94% of the variance explained, respectively) (**TABLE 4.2**). Larval dispersal potential was a significant factor consistently affecting log F_{ST} values from the substantial coverage dataset, but migratory habits was the variable that had the most consistent relationship with log F_{ST} values across all models and for both coverage datasets.

TABLE 4.2 Multiple regression models constructed to identify whether trait combinations affected population structure on the log transformed FST values. Linear models were run separately for studies covering different geographical coverage. LDP stand for Larval Dispersal Potential. If a level was significant for a given factor, that factor is highlighted in bold. Words that are in bold had p<0.05, and bold underlined had p<0.01. Best fit models can be evaluated based on the maximum adjusted r2, and on the lowest AIC and corrected AIC values.

Coverage	Linear Model	Variance explained (%)	Adjusted r^2	AIC	AICc
Substantial coverage	**LDP**	23	0.19	70.65	72.25
Substantial coverage	**LDP** + Habitat	23	0.08	74.61	79.22
Substantial coverage	**LDP** + Habitat + **Migration**	61	0.41	67.90	82.30
Substantial coverage	**LDP** + **Habitat** + **Migration** + **Depth Zone**	75	0.54	63.63	91.13
Substantial coverage	**LDP** + Habitat + **Migration** + Depth Zone + Distribution	75	0.43	67.60	119.60
Entire coverage	LDP	0.009	-0.09	52.60	55.27
Entire coverage	LDP + Habitat	6	-0.25	55.72	64.30
Entire coverage	LDP + Habitat + Migration	52	0.04	53.12	89.12
Entire coverage	**LDP** + **Habitat** + **Migration** + **Depth Zone**	94	0.82	30.15	140.15
Entire coverage	LDP + **Habitat** + **Migration** + **Depth Zone** + Distribution	94	0.76	32.03	296.03

The interdependence of larval dispersal potential and migration pattern in affecting F_{ST} values is illustrated with (**FIGURE 4.5**) which highlights that some of the specimens with the lowest F_{ST} values amongst low larval dispersal species are also migratory or highly migratory species, whereas specimens with the highest F_{ST} values amongst high larval dispersal species happen to be species that display no or limited migratory behaviours.

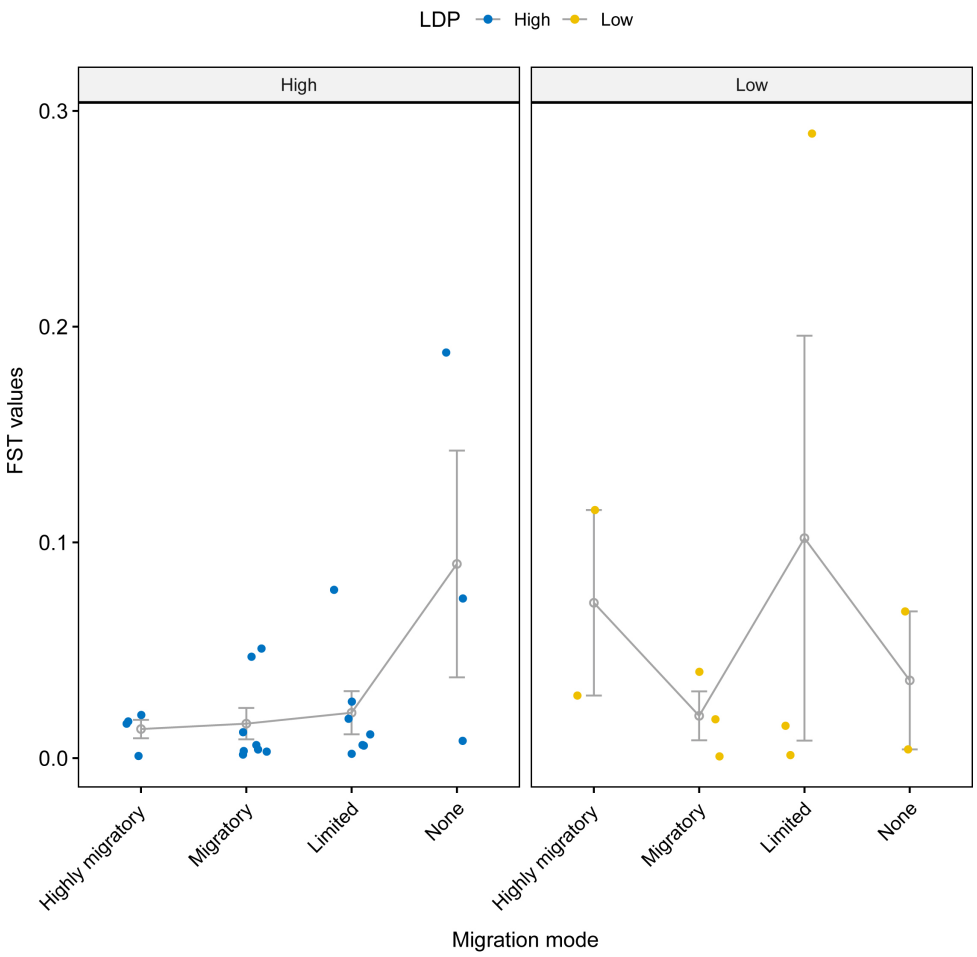


FIGURE 4.5 The effect of larval dispersal potential and migration on F_{ST} values.

4.5 Discussion

Both stable isotope and genetic methods effectively determine provenance in at least some species and in some areas (Carrera & Gallardo, 2017; FishPopTrace, 2013; Kim et al., 2015; Martinsohn et al., 2019; Nielsen, Cariani, et al., 2012; Rampazzo et al., 2020) yet, for both methods, only a limited number of applied verification studies have been carried out, often with some degree of dependency between samples used to define population characteristics and those used to estimate assignment accuracy. Here, we

introduce an operational framework to evaluate the species and areas for which it is reasonable to expect that genetic or stable isotope tools will prove useful. Biological traits proved relevant when evaluating species' population genetic structure and could therefore be used to inform when genetic provenance testing tools may be useful. Use of global mechanistic isoscape models appear beneficial in distinguishing which fishery areas are likely to be isotopically distinct, and therefore where isotopic provenance testing tools may be used. In addition, combining genetic and stable isotope tools can, in some cases, increase provenance verification power and decrease misassignment errors. The guidance provided here can also serve the important role of identifying those species or stocks where neither genetic nor isotopic approaches are expected to resolve provenance so that effort is spent on alternative solutions to provenance verification.

Genetic tools

Though it was not possible to estimate the number of species for which genetic analysis would be expected to confidently determine point-of-origin, we identified some traits which seem to affect population genetic structure and should be accounted for in provenance testing. Larval dispersal potential and migration mode could be used to guide expectation to find population genetic structure. Family might be a good predictor of population genetic structure, with some families exhibiting little range in structure levels and others exhibiting a larger range making structure potentially more difficult to predict. These results corroborate some of the observations reported by Bradbury et al., (2008) who noted that genetic differentiation calculated using F_{ST} values varied significantly across a range of higher-order taxa (e.g. polychaetes, crustaceans, echinoderms, teleosts, etc.). Irrespective of family however, identifying clusters of species based on the traits they share (**FIGURE 4.3**) might help decide whether genetic tools should be further investigated or not for any given species.

Larval dispersal potential was an important factor in determining population genetic structure which contrasts with a number of studies examining pelagic larval

duration as a factor (Galarza et al., 2009; Riginos et al., 2011; Weersing & Toonen, 2009). This discrepancy might be due to methodological differences in our assessment of larval dispersal potential compared to pelagic larval duration. Based on the assumption that larvae are planktonic and easily advected (Bohonak, 1999; Doherty et al., 1995; Waples, 1987) larval dispersal potential has previously been defined as the pelagic larval duration of a marine species (Weersing & Toonen, 2009). In an attempt to use a more comprehensive early life history trait however, we here not only accounted for larval pelagic duration information when available, but also used a thorough literature review to identify a variety of additional, non-exclusive set of variables such as larval type, larvae buoyancy and diel vertical migration behaviour, larval homing behaviour and swimming abilities, as well as observed or modelled advection patterns. Consequently, while our assessment of larval dispersal potential is a multidimensional, qualitative factor, it accounts for a more comprehensive description of early life-history.

When exploring whether methodological procedures could affect F_{ST} values and add to the variance observed within models we did not find any effect of marker type on F_{ST} values, which has previously been reported (Bradbury et al., 2008; Kinlan & Gaines, 2003) though none of the these studies included SNPs in their analysis. Although surprising, given the increasing use of outlying SNP loci for maximising genetic differences, the findings may simply reflect the general tendency to report and publish studies showing significant population structure far more frequently than those showing no spatial differentiation. On the other hand, geographical coverage of a given study did affect F_{ST} values, and similar metrics such as biogeography and study distance have also revealed such patterns (Riginos et al., 2011).

The meta-analysis for the DNA-based methods may suffer from biases and inaccuracy resulting from combining over 50 studies to draw inferences on population genetic structure. In the context of population genetic structure, we assume comparability of results between studies, but despite only selecting for studies using microsatellites or SNPs, the number of screened and selected loci as well as the genome coverage varied widely between studies thereby influencing robustness and comparability. Sampling

design, overall methodological approaches, and even species type will affect the precision and accuracy of the results and introduce important variances among studies, and we attempted to reduce some of the variance by accounting for marker type and geographical coverage.

When using life-history traits to predict the ability to detect geographical point-of-origin via population genetic structure, no one trait is sufficiently reliable (Bradbury et al., 2008; Galarza et al., 2009; Riginos et al., 2011). Reproductive isolation in marine organisms is not only dictated by a set of diverse species-specific ecological and life-history traits (Chopelet et al., 2009), but also by complex bathymetric and oceanographic conditions that differ between sites, seasons, and years (Selkoe et al., 2008). Despite these complexities, we were able to demonstrate that some traits, particularly larval dispersal potential and migratory habits, have the potential to affect population genetic structure and should be accounted for in studies attempting to use ecological traits to predict point-of-origin. Specifically, species with demonstrably low larval dispersal potential and limited migratory behaviour appear to exhibit the geographic structuring that would make an investment in genetic characterisation worthwhile. It should be noted however that the low F_{ST} values collected for our meta-analysis are not always synonymous with indiscernible population genetic structure and should not necessarily lead to an exclusion of genetics as an effective provenance testing tool. Rapidly developing genomic technologies will likely allow for detection of population genetic sub-structure in most scenarios (Barth et al., 2019; Bernatchez et al., 2017), as has been documented multiple times with commercially important species such as Atlantic cod (Barth et al., 2019; Johansen et al., 2020; Willette et al., 2014) and with the advent of genome wide assembly, the development of SNP arrays and increasingly performant computational methods. The financial and technical investment for this level of scrutiny however – and its practical application – will likely be outside the scope of routine traceability testing for the majority of commercial species. For the purpose of this study, however, and considering the number of studies accounted for in the analyses presented

here, F_{ST} does offer a good estimate of population genetic structure and can provide guidance for the development of operational tools in time- and budget-limited contexts.

Roadmap for method selection

Based on the results, we developed a decision tree approach to help guide users in deciding what tool to use when, and whether to invest in further baseline and validation efforts (**FIGURE 4.6**). For each new sample requiring provenance testing, we propose initially verifying the species using genetic barcoding to ensure species substitution has not occurred. The second critical step is to identify the labelled fishery location, and all other areas or interest or possible locations in which the fish could have been caught. If genetic or isotopic assignment studies have been carried out for the species of interest, covering all likely fishery areas, the genetic profile or stable isotopic ratio of the sample can be compared to existing forensic databases for that species and stated provenance can be either confirmed or disputed. However, to date, assignment studies have not been carried out for many species, and, of those, all likely fishery areas have not yet been compared. Therefore, the next stage of the process would be to determine if the species is listed within those presented here (Link to data on Dryad) and to determine whether the labelled fishery area is isotopically distinct from all other possible fishery areas based on global mechanistic isoscapes, and/or whether the species is likely to have strong population structure.

If the labelled fishery area can be distinguished from all other fishery areas with greater than 75% accuracy, a targeted stable isotope study is proposed where known origin samples of the species of interest are collected and measured across the species range. A threshold of 75% accuracy was determined based on the discrete assignment results of herring fishery areas, with three regions displaying assignment accuracy results of greater than 74% (Bekkevold et al., 2015; Nielsen, Cariani, et al., 2012), compared with successful unpublished results from a targeted stable isotope study of herring in the same regions. If a strong population structure is predicted, a targeted genetic assignment study

is proposed, where known origin samples are collected from all possible populations. For both methods, known origin test samples should then be measured for their isotopic ratio or genetic profile and assigned to each discrete area or population to measure assignment accuracy. If assignment accuracy to the labelled region is greater than 90%, a full provenance test is proposed. If assignment accuracy is less than 90% for any one method, and the alternative method has also been explored, either a combined genetic and isotopic assignment method or an alternative method is proposed. For application of these techniques into real world scenarios, we recommend that the threshold values be adapted to suit the accuracy requirements of the specific question.

4.6 Conclusions

A stark conclusion from this study is the fact that there are very few genetic or isotopic assignment studies available. Exploring the level of population genetic structure between putative populations and determining the range and variation in isotopic ratios between individuals caught within different fishery areas are necessary steps towards evaluating the point-of-origin of a product. However very few studies progress toward estimating the probability with which specimens can be assigned back to their population of origin (Bekkevold et al., 2015; Drinan et al., 2018; FishPopTrace, 2013; Nielsen, Cariani, et al., 2012; C. N. Trueman et al., 2017; Zhang et al., 2019). Yet certainty around assignment is ultimately what stakeholders and management agencies are seeking. Whereas F_{ST} values can appear vague and variable, and the resolution of global predictive isoscapes can be too coarse, assignment probabilities to reference data on known origin can offer concrete evidence for supply chain provenance verification, especially when these might be associated with court cases and loss of accreditation.

Future users of provenance verification tools, including seafood processors, retailers, government enforcement agencies, and certification bodies will need to invest in bringing these tools to operational readiness and the framework developed here is a first step towards prioritising these efforts. The framework helps to discern whether genetic or

isotope tools might be successful in the application at hand, and the option to use a combined approach is also available. However, the increase in verification assurance needs to outweigh the costs involved with carrying out analyses for both genetic and stable isotope markers. Such cost will vary substantially, depending on the organisations faced with them; thus, future multi-stakeholder engagement would be desirable to identify the best strategies to meet such investments, which, in a majority of cases, offer a hardly replaceable step to achieve stock traceability.

While sustainability and ocean conservation concerns rise on international agendas (UN, 2019), even with the advancement of digital traceability tools, such as blockchain, independent verification via forensic tools will remain a crucial asset to provide assurance of provenance in global seafood supply chains. To ensure this goal can be realised at the scale needed to support global efforts for sustainable fisheries, advancing the issues presented here could help operationalize provenance testing, so that it will become more widespread, technically feasible and financially accessible.

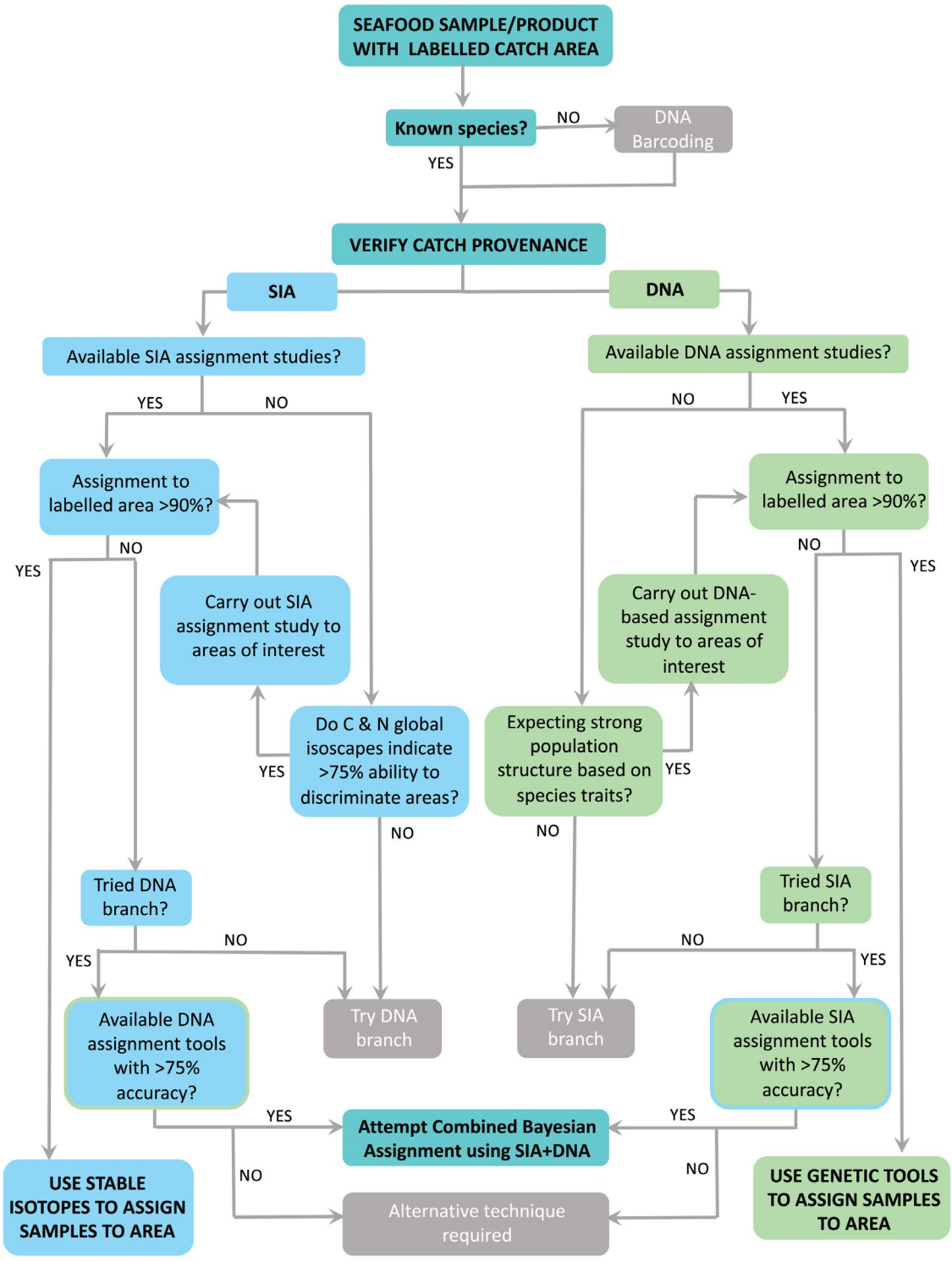


FIGURE 4.6 Decision tree depicting the recommended steps for selecting the most appropriate testing tool in verifying the provenance of a sample of fish or fish product, and providing considerations to help decide when DNA, stable isotope, or both markers combined are likely to be most effective.

Chapter 5
Below and beyond the species: Unravelling Atlantic cod catch location mislabelling across European markets

Authors: **Cusa, Marine**; Mariani, Stefano; Ogden, Rob; Baillie, Charles; Shum, Peter

5.1 Abstract

Over the last decade, the advent of affordable genetic tools has exposed seafood mislabelling worldwide. This has led to increased awareness and reduction of species mislabelling in some countries. On the other hand, the mislabelling of products' geographical origin has not yet received much attention as the methods to evaluate provenance are more complex than those required to identify species. Increasingly however, the use of DNA based tools for determining geographical point-of-origin of seafood are getting more accessible and emerging as a viable option. We conducted an international geographical point-of-origin market study, sampling Atlantic cod (*Gadus morhua*) from fish mongers and grocery stores across four European countries. Our goal was to identify whether there are any discrepancies between the catch areas indicated on the labels and the actual provenance of the product. We used a novel approach to target a set of nine previously identified diagnostic single nucleotide polymorphisms and followed an Illumina indexed sequencing workflow to multiplex a total of 126 market samples and 30 voucher specimens from known locations in a single sequencing run. We were able to identify whether the Atlantic cod product originated from either one of the following sources (1) the Northeast Arctic (Norway, Barents Sea, Bear Island, and Svalbard) or (2) the North Sea. Despite a confirmed low rate of species mislabelling, our study revealed that about 50% of the products were likely mislabelled with respect to their geographical point-of-origin. We further show that the rate of provenance mislabelling does not only differ among European countries, but also among retailer types. Beyond the established tools employed for species identification, DNA technologies devised for population assignment can help enforcement entities and environmental organisations monitor and prevent fraud related to geographical catch location.

Keywords

Authentication · Seafood fraud · Geographical origin · Traceability · Validation

5.2 Introduction

In much of modern Western fisheries management, sustainable stock harvest relies on knowledge collected during fisheries surveys and on the resulting establishment of quotas and potential restrictions. In the North East Atlantic, the total allowable catches (TACs) define the quantity of fish that can be harvested for a given species and given stock from year to year (Hilborn & Hilborn, 2019). In order for such regulations to be truly effective, stock management boundaries defined for commercial purposes must match biological population boundaries (Reiss et al., 2009). In Europe, stock assessment and management are largely based on a set of areas and subareas identified by the International Council for the Exploration of the Sea (ICES) (**FIGURE 5.1**).

Despite the lack of perfect concordance between management units and biologically relevant divisions (Reiss et al., 2009), this top-down approach to fisheries management has proved powerful in many cases and remains the framework under which modern Western fisheries are managed. The success of this approach for large scale fisheries relies heavily on scientific knowledge and on the understanding of stock dynamics, as well as on strong governance and implementation, and on fishermen's incentives to follow regulations. Illegal fishing practices and fraud or carelessness along the supply chain are an impediment to the successful application of these important measures. Molecular biology tools have revealed mendacious practices in the seafood industry, where the substitution of species is a widespread phenomenon across world markets. Much more inconspicuous and difficult to test is the mislabelling of specimen's catch location. The pervasiveness of this phenomenon in the seafood industry and its impact on fisheries sustainability remain largely unexplored, due to the paucity of provenance testing market studies conducted so far.

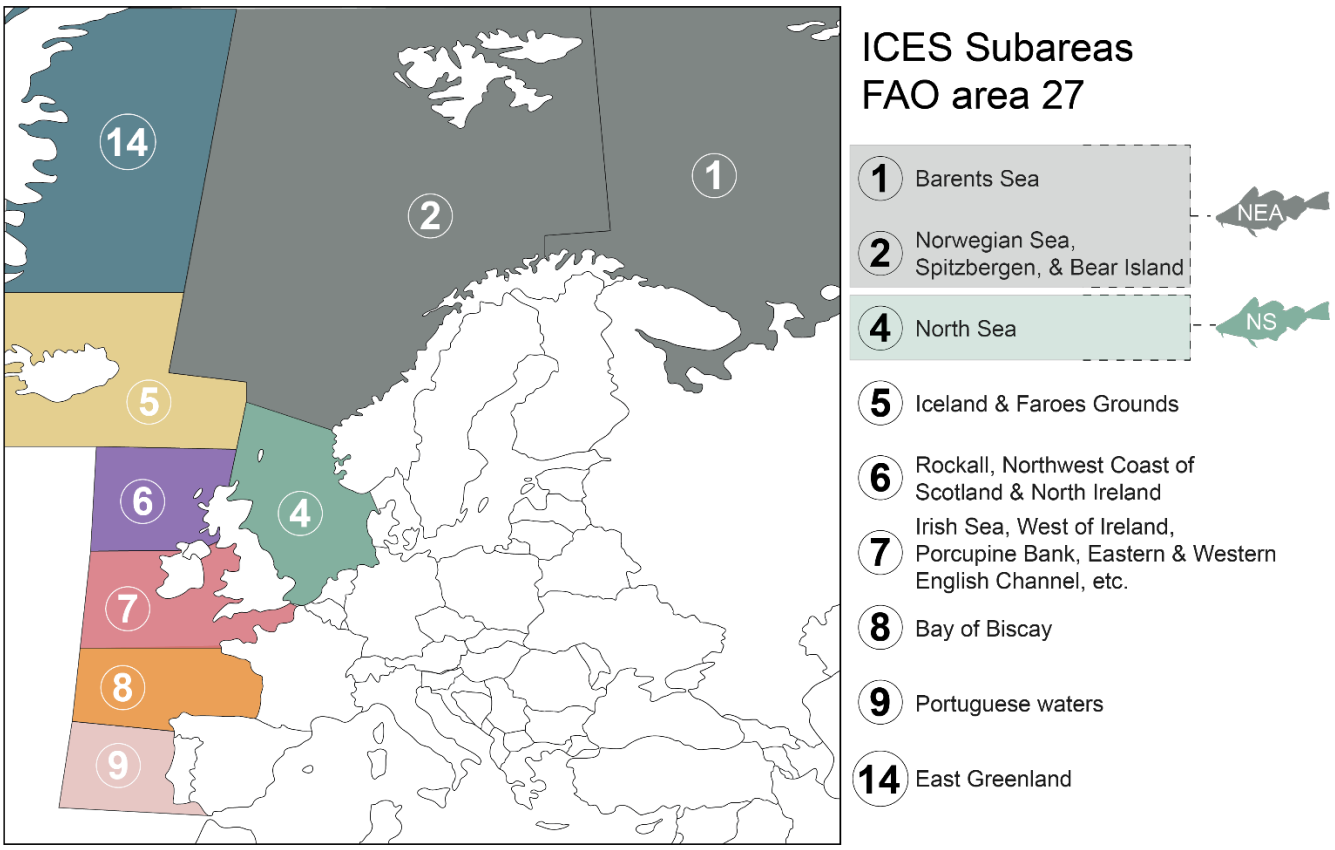


FIGURE 5.1 Map of the ICES subareas with highlights identifying the two regions and populations for which the Atlantic cod diagnostic SNPs were identified; the North East Arctic (NEA) encompasses ICES subarea 27.1 and 27.2 and is pictured in grey, and the North Sea (NS) encompasses ICES subarea 27.4 and is pictured in green. The other subareas all depict zones in which at least one marketed cod specimen was claimed to have been fished.

Atlantic cod is distributed throughout most of the North Atlantic and as far north as Svalbard and North Greenland (**FIGURE 5.2**). It has a long history of harvest and consumption and its interplay with human societies over many centuries is possibly unparalleled (Kurlansky, 1997). Presently, Norway and Iceland are the main harvesters of Atlantic cod, and many European countries have dedicated cod fishing fleets equipped for long fishing trips to the Norwegian sea, Barents Sea, Baltic Sea, and North Sea. Atlantic cod is a staple dish in many countries, such as klippfisk in Norway and Iceland, fish & chips in the UK, or Bacalhau in Portugal. Despite what appeared to be some promising stock recoveries, the Atlantic cod stocks of the North Sea and Baltic Sea recently underwent sudden and substantial depletion, leading to severe fishing restrictions and a zero catch advised for 2020 (ICES, 2019). Following this advice, certification bodies, such

as the Marine Stewardship Council (MSC) have promptly suspended their certification of the North Sea cod fisheries (MSC, 2019).

Both the economical importance of this fish and the stock collapse of the Atlantic Northwest fishery have led an important body of research to be conducted on Atlantic cod, including a number of population genetic studies. As a result, there are few commercial fish species as well studied as Atlantic cod. As early as 1965, Knud Sick attempted to evaluate the boundaries of cod populations in the Baltic Sea and Danish straits using genetic methods (Sick, 1965). It is now fairly well established that Atlantic cod demonstrates varied levels of population structure throughout the Atlantic with some populations being quite distinct from others. Microsatellites and nuclear DNA restriction fragment length polymorphism have demonstrated that the Northeast and Northwest populations can be easily differentiated as well as the Barents sea population from other Northeast Atlantic populations (Hutchinson et al., 2001; Jónsdóttir et al., 2003; O'Leary et al., 2007; Skarstein et al., 2007). Studies have also noted population structure among Northeast Atlantic populations (Nielsen et al., 2001; Pogson & Fevolden, 2003; Sarvas & Fevolden, 2005) and more recently, finer resolution of Northeast Atlantic cod population structure was obtained using Single Nucleotide Polymorphisms (SNPs) easily differentiating between North East Arctic, North Sea, and Baltic Sea populations (Nielsen et al., 2009; Poulsen et al., 2011).

The advance in genetic methods has permitted the exploration of forensic studies in the seafood industry and mislabelling of Atlantic cod in the European market was first reported using the Cytochrome c Oxidase barcode by Miller & Mariani in 2010 (Miller & Mariani, 2010). Since then, many studies have denounced the mislabelling of fish species in markets and restaurants throughout Europe, leading to more rigorous supply chain regulations. The exposed controversy around seafood mislabelling has spurred a flurry of reactions from the media and general public resulting in a lower rate of species mislabelling in Europe (Mariani et al., 2015) though resurgence of this phenomenon have recently been observed in some EU countries (Feldmann et al., 2021). Following the

denouncement of species mislabelling in the seafood industry, DNA-based methods have also been suggested as a potential tool for catching fraudulent behavior regarding product geographical catch location (Martinsohn et al., 2019; Ogden & Linacre, 2015). In terms of methodologies, this is a more complex endeavour, requiring more specialized tools (Cusa et al., 2021 - See Appendix). This has resulted in the slow development of such tools despite the economical benefits it could bring if used for law enforcement purposes (Martinsohn et al., 2019; Nielsen, Hemmer-Hansen, et al., 2012). As a result, the mislabelling of catch location on EU seafood marketed products remained, up to this point, unexplored.

Following the EU 7th Framework Programme FishPopTrace (FPT) project which was aimed at developing tools to evaluate the geographical point-of-origin of commercial fish species, members of the TRACE wildlife forensics network developed a Standard Operating Procedure (SOP) for the Department for Environment, Food & Rural Affairs (DEFRA) in the UK to provide a simplified method for the geographical assignment of Atlantic cod, Hake (*Merluccius merluccius*), Sole (*Solea solea*), and Herring (*Clupea harengus*) (Ogden & Murray-Dickson, 2014). The team performed genome-wide screening on hundreds of reference specimens to find SNP panels for each species and isolate the top-ranked SNPs for geographic assignment. A SOP was then developed via which interested parties could easily screen for these diagnostic SNPs and assign samples back to their population of origin. The authors settled for an economically advantageous KASP chemistry assay which proved difficult to replicate as it lacked in rigour and performed poorly for some species. They concluded that further optimisation of the KASP assay was required if it was to be used as a SOP. Using a total of 942 reference cod genotype data from known geographical locations, they screened through 1290 SNPs to identify a SNP-marker panel that could discriminate between sets of two populations, the North East Arctic (NEA) and the North Sea (NS) populations. The authors narrowed down the panel to 9 highly diagnostic SNPs that could discriminate between individuals from these two cod populations with 98% certainty. They further discuss that the 2% misassignment observed across the two populations might be attributed to the KASP chemistry assay which

required extensive optimization and still did not perform as desired.

It is the context of the recent cod stock decline in the Northeast Atlantic – as well as the various stock-specific regulations imposed by the EU – that we explored new ways of testing the provenance of the Northeast Atlantic cod using the same set of 9 diagnostic SNPs. If regulations are to effectively help the rebuilding of cod stocks, it has now more than ever become essential to catch any potential fraud that could impede on this ambitious and essential goal. Indeed, if North Sea marketed cod is sold as Barents Sea cod, this would indicate a traceability failure along the supply chain and would prevent mindful consumers from selecting the more sustainably sourced product. In this study, we developed a parallel sequencing protocol to segregate between the NEA and the NS cod populations using the 9 diagnostic SNPs identified by Ogden and Murray-Dickson (2014), and tested the technique on cod samples of known origin to evaluate the efficacy and accuracy of the method in assigning specimens back to their population of origin. We then ran the first international cod provenance market study using these genetic tools on cod samples from both fish mongers and supermarkets in the United-Kingdom, France, Germany, and Spain. The aim of this study was three-fold **(1)** to develop a parallel sequencing method that reduces costs by allowing to genotype many samples simultaneously for testing geographical provenance of marketed cod, **(2)** to test whether marketed cod match their claimed area of origin by testing specimens using a North East Arctic vs North Sea assay, and **(3)** to evaluate some of the factors that might affect mislabelling rate of cod catch area.

5.3 Methods

5.3.1 Marketed cod sampling

Sampling was conducted in various cities and coastal communities around the United-Kingdom, France, Germany, and Spain (**FIGURE 5.2**). We attempted to get a relatively equal number of samples from fish mongers and from supermarkets though that was not always

possible. In the UK, a total of 47 cod were purchased from towns and cities in the northern part of the country (Manchester, Liverpool, Newcastle, Fraserburgh, Peterhead, Gourdon, Arbroath, and Edinburgh). In France, a total of 28 cod were purchased in various supermarkets and fish mongers from Paris and from towns in the region of Brittany (St. Malo, Guilvinec, Concarneau, and Quimper). In Germany 10 cod were sampled from Hamburg and another 10 from Kiel, and in Spain 13 cod were sampled from Vigo (**Table S5.1**). Additionally, several samples of known origin (positive controls) were received from partner institutions to validate the method used for geographical assignment: 10 cod from the East Barents Sea received from the University of Oslo, 10 cod from West Svalbard received from the Norwegian Polar Institute in Tromsø, and 10 cod from the North Sea received from the Centre for Environment, Fisheries and Aquaculture Science in the UK. Finally, another 22 UK-sampled cod from the 2013 Labelfish project were included in the study (Mariani et al. 2015). Within given supermarkets, a maximum of three packages with the same EU code were purchased with the intent to diversify the brands and processing factories as much as possible (Miller, Jessel, et al., 2012). Pictures of labelled packages were taken, FAO area and ICES subareas as indicated on the packages were recorded for each sample, and in the case of fish mongers, if labels were not present on the stall, the sellers verbally confirmed the catch location of the Atlantic cod they sold. Samples were stored in a tube with silica beads or in ethanol.

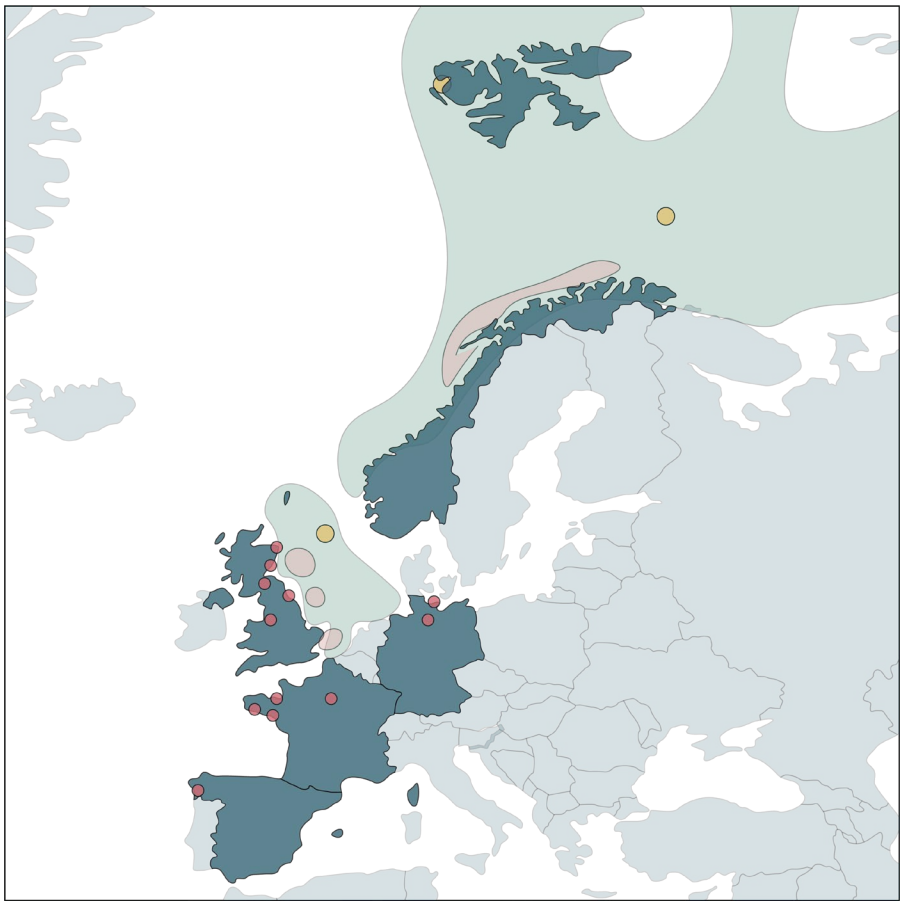


FIGURE 5.2 Map illustrating the various locations where Atlantic cod specimens were sampled for **a)** the market study (pink dots), and **b)** the validation of the method (yellow dots). The distribution of the two populations of interest (North East Arctic & North Sea) is highlighted in pink for the spawning grounds and green for the adults.

5.3.2 Compliance with EU labelling rules on ICES subareas

We evaluated compliance with the EU regulation on seafood labels by recording the number of ICES subareas indicated for each product. According to Regulation (EU) No 1379/2013 and Regulation (EU) 1169/2011, it is mandatory for any EU prepacked and non-prepacked fresh or frozen seafood product to not only state the scientific name of the species sold, but also to display detailed information on the catch area for fish caught at sea. For any fish that was caught in the Northeast Atlantic, the ICES subarea or division must be indicated and defined clearly so that consumers can understand the provenance of the seafood. The ICES Northeast Atlantic subareas relevant for this study within FAO area 27 are unambiguously defined on the FAO resource website and are as follows:

Barents Sea (Subarea 27.1), Norwegian Sea, Spitzbergen, and Bear Island (Subarea 27.2), Skagerrak, Kattegat, Sound, Belt Sea, and Baltic Sea (Subarea 27.3), North Sea (Subarea 27.4), Iceland and Faroes Grounds (Subarea 27.5), Rockall, Northwest Coast of Scotland and North Ireland (Subarea 27.6), Irish Sea, West of Ireland, Porcupine Bank, Eastern and Western English Channel, etc (Subarea 27.7), Bay of Biscay (Subarea 27.8), East Greenland (Subarea 27.14) (**FIGURE 5.1**).

5.3.3 Samples DNA extractions

Silica beads were removed from tissue samples or samples were removed from ethanol and tabbed on tissue paper and a biopsy was taken for the DNA extraction. The DNA extraction was conducted for all 28 French samples using the E.Z.N.A. tissue DNA Kit from omega bio-tek, Inc. The rest of the extractions for the UK, Spanish, German, Norwegian, and LabelFish samples were conducted using a Mu-DNA extraction protocol (Sellers et al., 2018) that proved more economically viable. The quality and quantity of DNA was evaluated both with a Qbit and a nanodrop.

5.3.4 Species identification and validation

Identification and validation of species were conducted in-house using the FASTFISH-ID™ protocol and reagents (Naaum et al., 2021 - See Appendix). A qPCR master mix was prepared by adding:

10.25 µL ((# of samples * 1.5) + 1) Fluorescent FASTFISH-ID™ reagents
0.25 µL ((# of samples * 1.5) + 1) Taq polymeraze

In each PCR tube, 11.34 µL of master mix was added to 2.16 µL of DNA. The FASTFISH-ID™ assay was then run in the Magnetic Induction Cycler for two hours with a thermal profile described in Naaum et al. (2021) and the results were imported onto the FASTFISH-ID™ online software for species identification.

5.3.5 Generating the primers for parallel sequencing

The sequences containing the nine diagnostic SNPs identified by (Ogden & Murray-Dickson, 2014) are about 120 nucleotide long and the primers were designed using an online primer design software (Primer3.ut.ee) and chosen to require a melting temperature (TM) of 60°C ± 2°C in an attempt to avoid multiple PCR profiles for the amplification of different segments. The 5' end primers were left as suggested by the primer design software whereas the 3' end primers were reversed and complementary from the primers designed by the software. Based on the protocol from Meyer and Kircher (2010) a 33-34 nucleotides long Illumina sequencing primer tail was added to the 5' end of the forward and reverse amplicon primers (**FIGURE 5.3**). Meyer and Kircher (2010) describe how to design the indexed barcodes containing both the index that will allow to multiplex each sample, and the Illumina flow cell binding sequences, called P5 and P7. The index sequences themselves are designed using a python script that was written by the bioinformatics group of the Max Planck Institute for Evolutionary Anthropology (https://bioinf.eva.mpg.de/multiplex/). Both tailed and non-tailed forward and reverse primers were then ordered as customized 5' → 3' DNA oligos from Eurofins for all nine amplicons (**TABLE 5.1**).

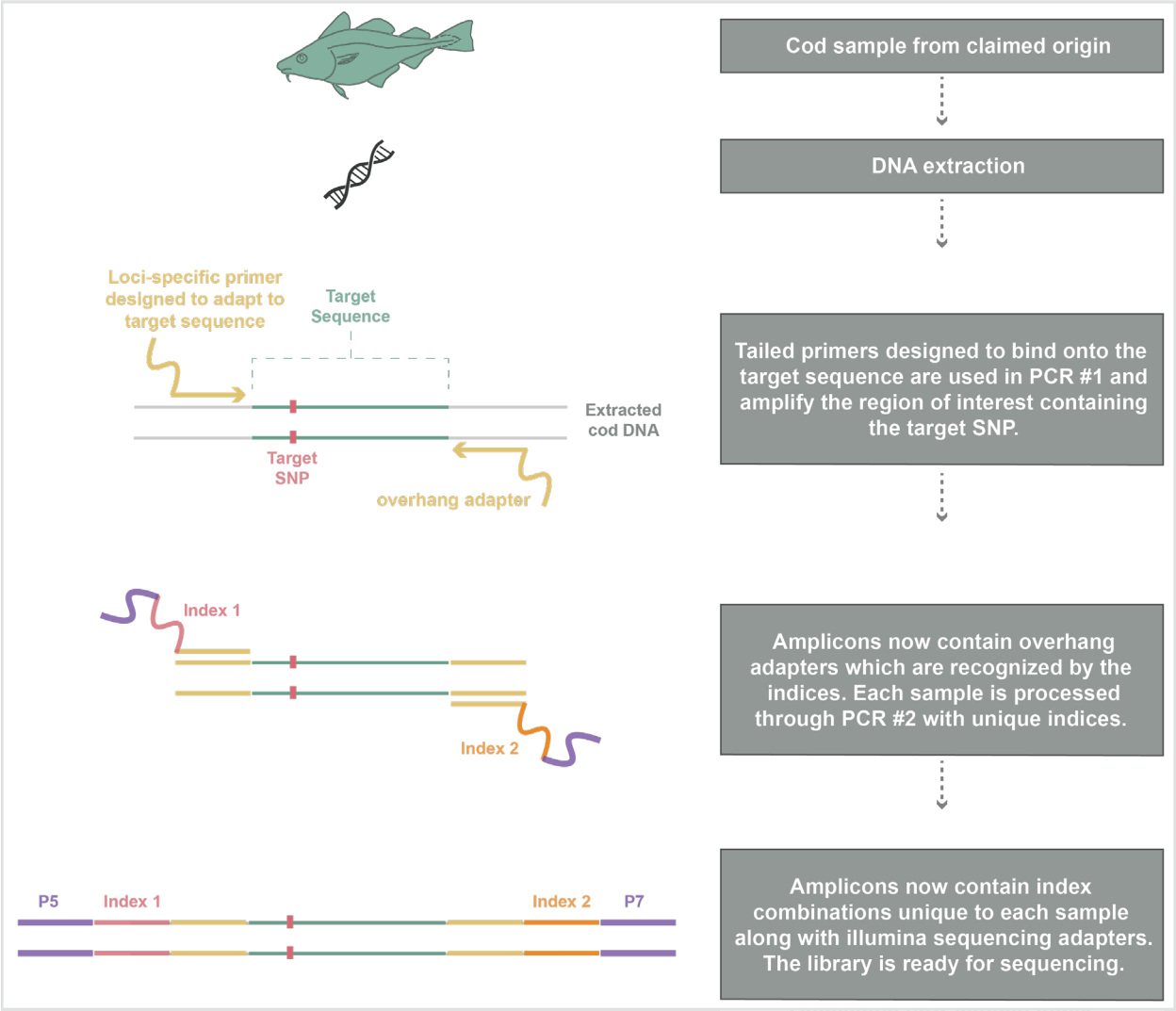


FIGURE 5.3 An illustration highlighting the SNP genotyping process using a two-step PCR and unique indices allowing to multiplex all samples in a single parallel sequencing run. The P5 and P7 are flow cell binding sequences required for the Illumina sequencing technology.

TABLE 5.1 a) Primer tail oligonucleotide sequence added to the 5' end of the forward and reverse amplicon primers, **b)** gene names that contain each of the nine targeted SNPs and their associated primers, **c)** target amplicon containing the primer binding site in red, the SNP site in green (marked as a bold 'N'), and the melting temperatures for forward and reverse primers.

a)

Oligo ID	Sequence
P5 adapter	ACACTCTTTCCCTACACGACGCTCTTCCGATCT + 5'Forward Primer3'
P7 adapter	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT + 5'Reverse&Complementary Primer3'

b)

Name	Left-Forward	Right-Reverse	Tm-Right	Tm-Left	Tm profile
1011C1CO1.158	AGTGAGTGCTGCAGAGAACC	GCTAACGGCTAGCCTCTGAT	59.3 C	60.0 C	High
1612C1CO1.537	ACTAACCCGTTACCATGCCA	TGGTATCCTTGAAGTGGACTTGT	59.3 C	59.0 C	High
1637C1CO2.549	AGGTGACAAAAATTAAGCTCATGAGA	GCCCGGAGAAGCCATTTGAA	61.0 C	59.3 C	High
1beta787	TGGGTTGTGTCTGCGCAG	TTTTTCTCGACACACAAGCGC	60.3 C	60.6 C	High
2193C1CO1.209	CTCAAGGCTCCCAGGAATGG	GTGAGGCTACAGACCGGC	59.8 C	60.1 C	High
2276C1CO1.375	ACAAATTTGCGAAACATGCA	TTTATTACAAGTATCCTCTTTTACGAG	55.6 C	55.7 C	Low
4529C2CO1.276	ACTGTGTGAAATGAGTCAACCCT	GTCAATTTGGGATTTAGCCATCGT	59.9 C	60.1 C	High
6375C2CO1.601	GAGTCAAATCCATCTGCATTAATAGA	AGCAAACTTTCATCTGGGCA	58.0 C	57.4 C	High
8727C1CO1.443	AGATTGTACTTACATACCATCCACA	CCACTTTCGTAATGACATGGGC	59.9 C	57.6 C	Low

c)

Amplicon
AGTGAGTGCTGCAGAGAACCAATGCAAAACTAAACTGAAACAAGTTAGCATNAGCATCTCTGGACTAAAGCGATGGCTTTAATCAGAGGCTAGCCGTTAGC
ACTAACCCGTTACCATGCCAAMCTACCATTCAGGGGGTGGTGGGGGCANGTATCATGGCGGTGGAGAGCTACCCTGATGACAAGTCCACTTCAAGGATACCA
AGGTGACAAAAATTAAGCTCATGAGAAAAAGTGGACATTAACAGCAAACAAATGTNCAAAAAGCAGTCAATGGTTTTGAGGCCTCTTTCAAATGGCTTCTCCGGGC
TGGGTTGTGTCTGCGCAGGTGCGTGGAGACAAACCTACATGTTTATGCGTGYATGCGAGTNGGCCTATATGCTTGCAGGTGCGCTTGTGTGTCGAGAAAAA
CTCAAGGCTCCCAGGAATGGAGAAACAGCCAGCCCGACAGCTGATGTAAAGCCAGNAGTTCCTTAGTGCTTTCTCCATCGAGGCCAGGTCAACTCGGGGCCGGTCTGTAGCCTCAC
ACAAATTTGCGAAACATGCAAAGATCTTTGAGAATGCTTCACAAATACCAAAAATGTAAGNAACATTGTTTCTACTCGTAAAAGAGGATACTTGTAATAAA
ACTGTGTGAAATGAGTCAACCCTTTAGACCTTCACTAATGAAATTAAATAATTCAGCCANGCATCTTCAGTCGCAGGGCCATACGATGGCTAAATCCCAAATTGAC
GAGTCAAATCCATCTGCATTAATAGATTTCAAACAATGAATCAACTTAGCTTTCGGSGATNGAGTGCACCACAAACTGCAAACATTATGCCCAGATGAAAGTTTGCT
AGATTGTACTTACATACCATCCACATAGTAAACAATAAATATAACAAACTAACTAAAATCNTACTGTTTTAATTGTGACAGCCCATGTCATTACGAAAGTGG

5.3.6 Optimization of primer assay and PCR1 for parallel sequencing

The optimization of the parallel sequencing protocol was performed using a subset of 28 samples. Adequate PCR profile temperatures needed to be established for each modified primer pair, with the consideration that tailed primers might not respond to the same temperature profile as non-tailed primers. We first trialed PCR profiles based on the New England BioLabs, Inc. (NEB) recommendations for the Taq 2X Master Mix with an annealing temperature of 58°C roughly, 1-3°C lowed than the primers TM. The run failed to amplify the target amplicons which prompted us to test various annealing temperatures, concentrations of the PCR library preparation, and a number of cycles before finally settling on an optimized assay and run profile.

We then amplified each of the nine primers individually for all remaining samples in 10 µL reactions containing 5 µL of Taq 2X Master Mix from NEB, 0.5 µL of the combined forward and reversed primers (with each primer pair in separate reaction), 0.15 µL of BSA, 2 µL of DNA (with each sample in separate reaction), and 2.35 µL of molecular grade water. We set the annealing temperature to 56°C for high TM primers and 45°C for low TM primers (**TABLE 5.1 b**) and followed an optimized PCR profile with an initial activation phase at 95°C for 30 seconds, followed by 30 cycles of denaturation at 95°C for 30 seconds, annealing at 54°C or 45°C for 30 seconds, extension at 72°C for 40 seconds, and finishing with a single final extension step at 72°C for 5 minutes. The quality and integrity of the PCR product was then checked using gel electrophoresis and any failed reaction was attempted again, or in some circumstances, required to conduct new DNA extractions before re-running the PCR.

5.3.7 Attachment of indices and sequencing adapters during PCR2 for parallel sequencing

Once the best primer runs were selected for each sample, all nine SNP amplicons were combined into a single solution for each sample. The combined product of PCR1 for each sample was then cleaned using the Qiagen Inc, MinElute PCR Purification Kit to remove

primer-dimers. 163 unique combinations of P5 and P7 barcodes (i.e. indices) were generated. A second PCR was then ran for each sample in duplicate to permit the annealing of the unique barcodes to each sample amplicons. The PCR2 product was then cleaned using the Qiagen Inc, MinElute PCR Purification Kit to remove anything smaller than the target amplicon. A gel was run to visualise the successful amplification of the indexed amplicons for each sample.

5.3.8 Amplicon library preparation for parallel sequencing; denaturation and loading

The cleaned DNA product from PCR2 was quantified for each sample using a Qbit in order to evaluate the volume of each sample needed for the final multiplexed solution. Following these volume calculations, all samples were multiplexed into a single solution. The DNA from the multiplexed solution was then quantified once more using the Rotorgene qPCR, the Qbit and the Tape Station to evaluate the dilution required for sequencing. 1.2 µL of library was added to 108.8 µL of EB buffer, and the DNA from the solution was quantified again using the Qbit resulting in a 0.5467 ng/µL solution. A 4 nM library was then prepared. To verify that the concentration was adequate we calculated it with the following equation: (0.5467 * 1,000,000)/128,700 = 4.25 nM. This resulted in 5 µL pool library (4nM) + 5 µL Sodium Hydroxide (NaOH) = 2.12 nM. To stop the denaturation step, we added 990 µL of HT1 hybridization buffer, diluting the library further by 1000-fold, and leaving us with about 21.2 pM. We wanted to load 8 pM for a final desired volume of 600 µL onto the MiSeq cartridge. To do this, we mixed 226.42 µL of 19.7 pM library with 10 µL of 20 pM PhyX and with 363.58 µL HT1. The MiSeq v3 cartridge (150-cycle) was then loaded into the MiSeq and the sequencing run started.

5.3.9 Bioinformatic steps

The Illumina paired end sequences are reported as R1 (forward) and R2 (reverse) files. Given that different SNP markers have variable lengths, we first concatenated all the R1 and R2 samples into two separate R1 and R2 files. We then used the fastqc package

(Andrews, 2010) to check the quality of the sequences using a QC analysis. FastQC reads sequence files and generates a quality control report for each of them. The reports contain information that can help identify problems that may arise in the data. In the case of our data, the bases after 110 bp needed to be trimmed. We used the package Obicut (Boyer et al., 2016) to trim the data and trimmed each sequence object at 119bp. The package FLASH (Fast Length Adjustment of Short Reads) (Magoč & Salzberg, 2011) was then used to align and merge the reverse and forward sequences. We then used the GTseq pipeline (Campbell et al., 2015) to genotype each individual cod. The program finds the primers and search for a probe pattern, it then counts how many times each pattern is observed and calculates proportions of how many times each SNP is found. We were then able to visualize how many times each nucleotide was counted for the SNP of interest and evaluate if the individual loci were homozygous or heterozygous.

5.3.10 Data analysis and assignment of market samples

A DAPC, which is particularly well suited to identify genetic clusters as it maximizes the variance between groups and minimizes the variance within groups, was ran on the Atlantic cod reference specimens provided by the TRACE Wildlife Forensics Network and containing a total of 273 cod from the North Sea and 304 cod from the North East Arctic using the package 'adegenet' (Jombart, 2008) in R (R Core Team, 2019). Prior group (*k*) number was defined as *k*=2, and allele frequency from this dataset was used in the DAPC to describe genetic clusters using a single discriminant function.

To evaluate the position of our market and voucher samples relative to the reference samples, we first performed a Correspondence Analysis (CA) using the GENETIX 4.05.4 software (Belkhir et al., 2004). Genetic assignment analyses were then conducted using the GENECLASS2.0 software (Piry et al., 2004). We used the partial Bayesian approach developed by Rannala B & Mountain JL (1997) to evaluate population allele frequency based on the cod reference samples and assign market samples to those populations. We then ran an exclusion probability analysis using Monte-Carlo resampling with a simulated number of 10 000 individuals and a type one error set to 0.01. The voucher specimens

from known locations collected in 2020 (Svalbard, Barents Sea, and North Sea) were also included among the 'unknown location' samples to verify whether they would be attributed correctly back to their known population of origin. In order to meet forensic criteria that would be required in the court of law in case of mislabelling, we evaluated the exclusion probability associated with the most likely population and the likelihood that an individual would belong to the population it was assigned to, given the claim that it belongs to a different population. Therefore, for any individual for which the claimed catch location did not match the assignment location and for which these were either the NEA or the NS, we calculated the log-likelihood ratios (logLR) and evaluated the likelihood of the alternative hypothesis given the observed genotype.

To evaluate some of the factors that may have affected mislabelling rate of the point-of-origin of the market samples we performed a series of Pearson's Chi-squared tests using the 'gmodels' package (Warnes et al., 2018) and we conducted a Generalized Linear Model (glm) to evaluate the effect of MSC certification on the rate of mislabelling with retailer type as a confounding variable in R (R Core Team, 2019).

5.4 Results

5.4.1 Species identification

Species identification was performed in-house using the FASTFISH-ID™ technology developed by Thermagenix (Naaum et al., 2021). A minority of samples were identified either as *Gadus macrocephalus* (1 out of 133 market samples) and *Melanogrammus aeglefinus* (3 out of 133 market samples). All other samples were identified as *Gadus morhua* (121 out of 133 market samples), aside from 7 of them which were not identified, possibly due to poor DNA amplification or to sample degradation. This quick species identification step allowed us to evaluate within a few hours that out of 126 identifiable samples only 4 were mislabelled, bringing the species mislabelling rate to about 3%. All four mislabelled specimens were sampled in the UK, two of them came from

supermarkets and were sampled in 2013 and the other two came from fishmongers and were sampled in 2020.

5.4.2 Compliance with EU labelling rules on ICES subareas

Overall, and when excluding samples from 2013, 96% of all samples contained detailed information on the provenance of the product (i.e. ICES subareas or equivalent), with information being written on labels or communicated orally by retailers. This is a drastically different picture from the one observed in the UK in 2013 where a single sample out of 22 contained detailed information on the provenance of the product. Despite providing information on ICES subareas however, many packages still offered a relatively vague picture of the catch area of the product by indicating more than one ICES subarea. Fish mongers were more inclined to indicate a single subarea compared to supermarkets which sometimes included up to four or more subareas (**FIGURE 5.4**). Country of retail also affected the precision of the information communicated or found on the packages with Germany offering the most precise level of information (i.e. A single subarea indicated for 75% of the products) and the UK offering the least precise level of information (i.e. A single subarea for 57% of the products). Overall, indicating two subareas appeared a frequent occurrence in all sampled countries, particularly in packaged supermarket products, and labels from the UK were more prone to vagueness with up to three or four subareas being indicated for 22% of the products.

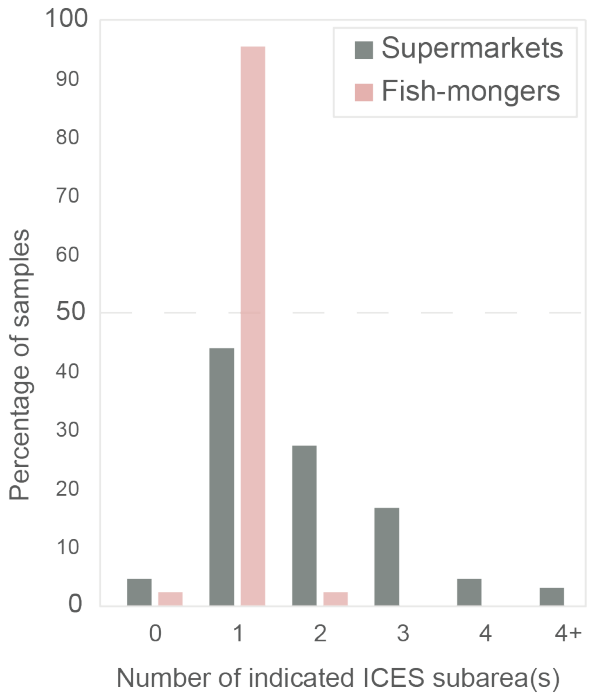


FIGURE 5.4 Bar-graph indicating the percentage of samples that contained information about ICES subareas on their package for all samples collected in 2020, '0' standing for 'no information available' and any other number indicating the number of sub-areas communicated or labelled on the package.

5.4.3 Parallel sequencing results and genetic assignment

A total of nine loci were used to verify the point of origin of 126 Atlantic cod samples purchased from supermarkets and fish mongers in the United Kingdom, France, Germany, and Spain. Four of the market samples were dropped as they were not Atlantic cod, and three were not included due to failed PCR amplification. Out of the 126 samples that were sequenced using this protocol, all 9 SNPs amplified successfully for 95 samples, 8 SNPs amplified successfully for 25 samples, and less than 8 SNPs amplified successfully for the remaining 6 samples.

As illustrated in a DAPC scatter plot for k=2, the two reference populations did structure distinctly, though some overlap is observed (**FIGURE 5.5**) indicating that misassignment might be a possibility and therefore calling for a conservative interpretation of assignment results.

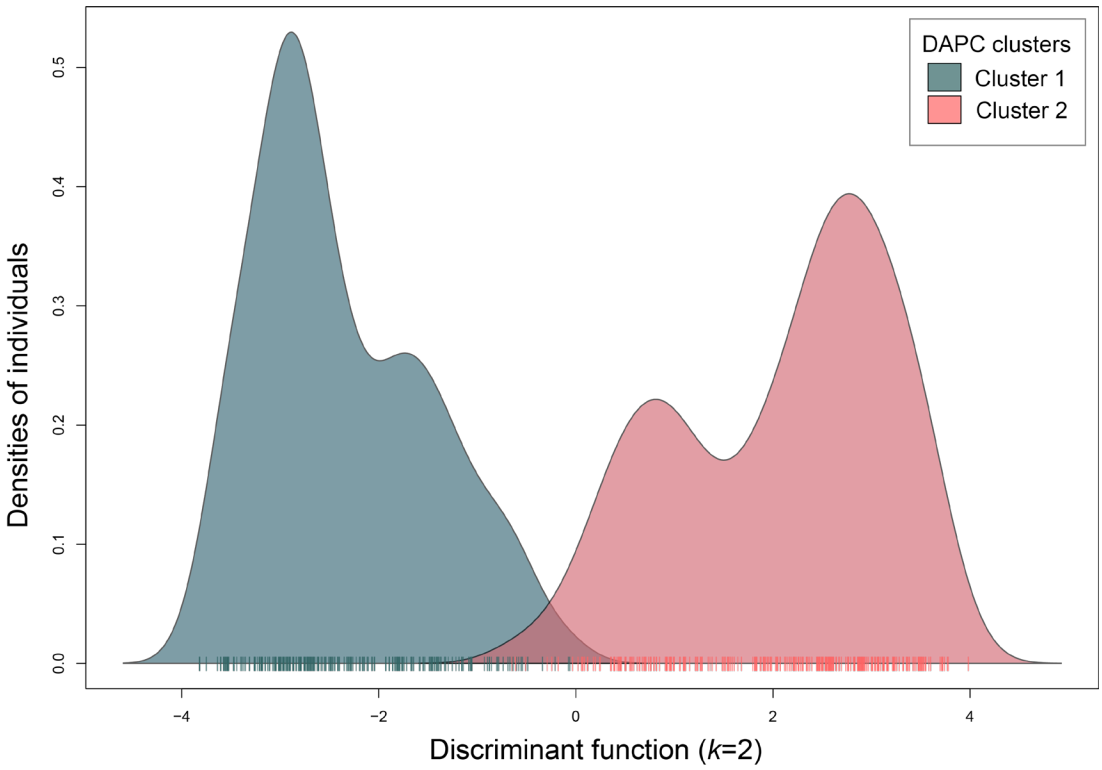


FIGURE 5.5 Plot illustrating densities of individuals from two different population clusters (*k*=2) as identified by a DAPC analysis on a single discriminant function.

Market samples and positive controls (voucher samples) were assigned to the reference populations using GENECLASS 2.0 and visualised using a CA, thereby forcing the samples in a two-dimensional frame, with Axis 1 explaining 26.60% of the variance observed and Axis 2 explaining 12.25% of the variance observed for a total of 38.84% of the variance being explained by these two dimensions (**FIGURE 5.6**). Two fairly distinct clusters can be observed corresponding to each reference population, NS and NEA, and voucher specimens (positive controls) and market samples are distributed both along Axis 1 and Axis 2 with North Sea voucher specimens clustering on the bottom right of the diagram and Barents Sea voucher specimens clustering in the top left of the diagram. Market samples fall somewhere within this diagonal with the majority of samples clustering in the top left indicating that most marketed specimens likely originated from the NEA (ICES subarea 27.1 and 27.2).

When using the Rannala and Mountain (1997) partial Bayesian model, assignment values from GENECLASS 2.0 ranged from 51.0% to 99.7% with a total of 83 out of 126 market samples being assigned with over 80% accuracy, and 54 samples being assigned with over 90% accuracy. The ten voucher specimens from the Barents Sea were assigned to the NEA with nine of them being assigned with >80% accuracy. Percent assignment accuracy was lower for specimens from Svalbard suggesting that they may have had a slightly different genetic makeup that prevented them from being reliably assigned. Svalbard specimens that were erroneously assigned to the NS either had low assignment accuracy (<75% assignment score) or low SNP amplification (<8 SNPs successfully amplified). Eight out of the ten voucher specimens from the North Sea were assigned to the NS with over 90% accuracy, one was assigned to the NS with over 70% accuracy, and one specimen was assigned to the NEA with a low assignment accuracy of 53.5%. The assignment of the voucher specimens indicated that a minimum of 8 amplified SNPs and 80% assignment score is likely to provide reliable assignment results for market samples. Overall, a total of 90 market samples were assigned to the NEA (63 samples with over 80% assignment score) and 36 market samples were assigned to the NS (20 samples with over 80% assignment score). This reflects the observations from the CA plot suggesting that the majority of samples originated from the NEA.

Using the logLR for mislabelled individuals, we evaluated that for all 13 individuals claimed to come from the NS and assigned to the NEA, the probability that they belonged to the NEA population was over 0.98. Similarly, for all 10 individuals claimed to come from the NEA and assigned to the NS, the probability that they belonged to the NS population was over 0.98. Additionally, 23 samples claimed to come from another region were assigned to the NEA and 12 were assigned to the NS. We could not calculate the logLR for these samples as we do not have genotype information for the claimed population. For these samples, we therefore relied solely on the assignment scores.

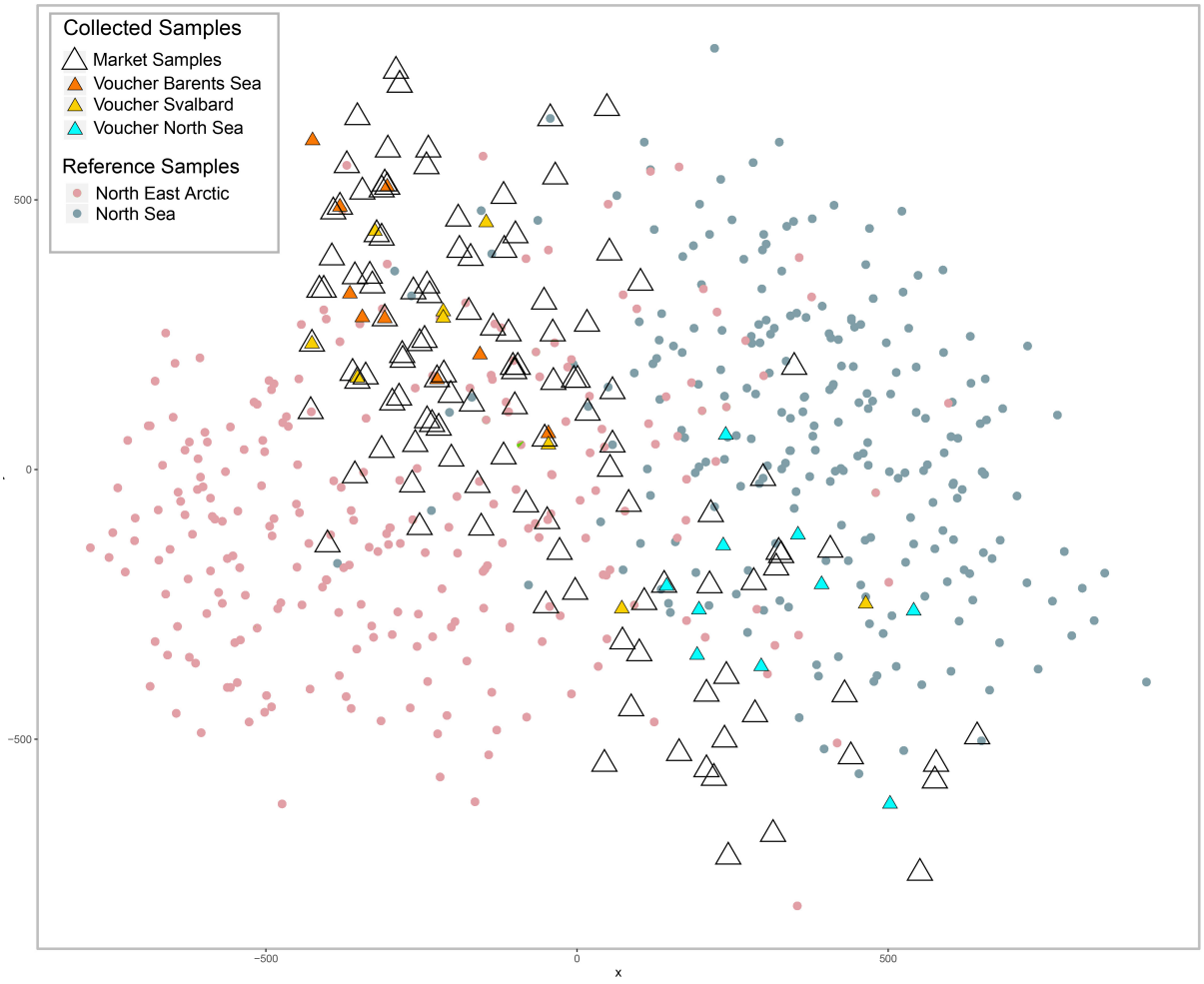


FIGURE 5.6 Correspondence Analysis illustrating the distribution of North East Arctic (pink dots) and North Sea (blue dots) reference samples along the two principal components that explain 38.84% o the variance. The market samples are represented with empty triangles and the voucher specimens (positive controls) with coloured triangles. The figure indicates that most market samples originate from the North East Arctic.

5.4.4 Point-of-origin discrepancies and mislabelling

The market sample analysis was performed using two assignment thresholds: a highly conservative >90%, as well as >80%, which, while being more relaxed, also prevents from "over-filtering" important and relevant data. Samples with lower assignment scores were removed from the analysis. Out of the market samples for which at least 8 SNPs had successfully amplified, 32% had an assignment score lower than 80% (33 out of 102 samples). Out of the voucher samples for which at least 8 SNPs had successfully amplified, 42% had an assignment score lower than 80% when including the samples

from Svalbard, and 15% when solely including the samples from the Barents Sea and from the North Sea (3 out of 20 samples).

When strictly selecting for samples that had an assignment score of 90% or 80% and for which ICES catch location information was provided, the overall mislabelling rate was 49% (22 mislabelled samples out of 45) and 50% (36 mislabelled samples out of 72) respectively. More specifically, 36% of fish mongers had provided accurate catch location information at 90% assignment accuracy and 26% at 80% assignment accuracy, whereas 69% of supermarket products were correctly labelled at 90% assignment accuracy and 64% at 80% assignment accuracy. This observed mislabelling rate differed significantly between fish mongers and supermarkets both for data containing assignment scores over 80% (X^2 (1, n = 72) = 10.01, p < 0.005) and for data containing assignment scores over 90% (X^2 (1, n = 45) = 8.09, p < 0.005). At 80% assignment accuracy, France displayed the highest level of mislabelling with only 25% of the samples being correctly labelled, followed by the UK with 54% of the samples being correctly labelled, and by Spain and Germany with 64% of the samples being correctly labelled. At 90% assignment accuracy, the picture changed slightly with both France and Spain displaying the highest level of mislabelling with only 33% of the samples being correctly labelled, followed by the UK and Germany with 63% and 67% of the samples being correctly labelled respectively (**FIGURE 5.7**) (**Table S5.1**).

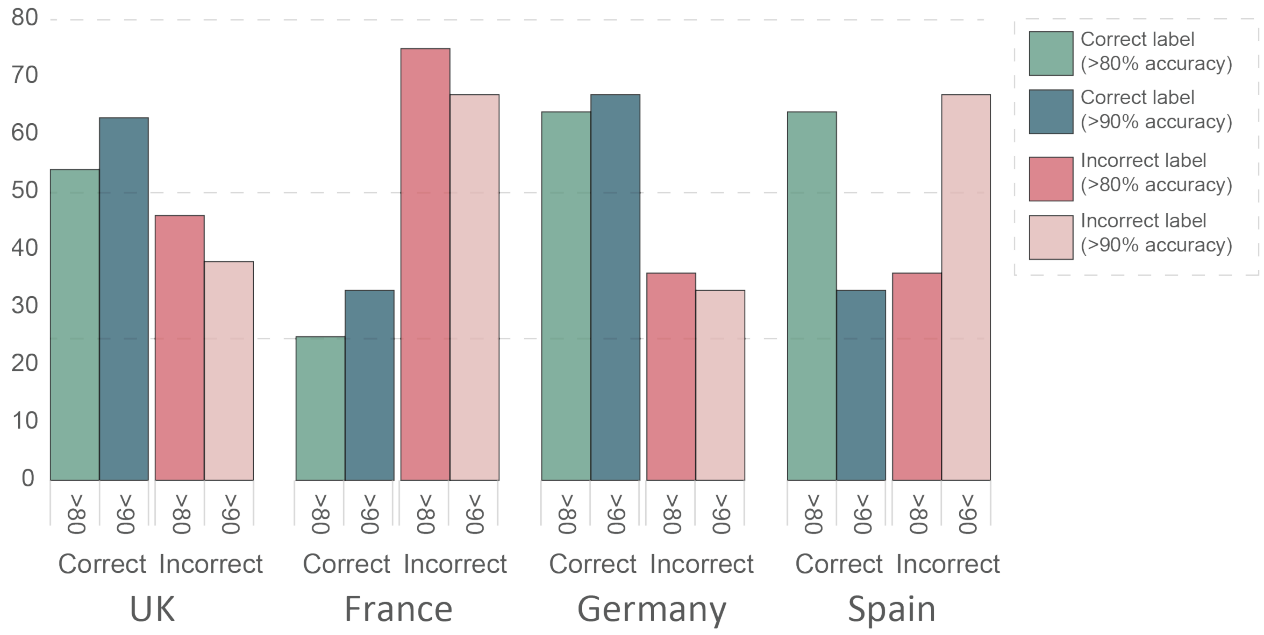


FIGURE 5.7 Mislabelling rate illustrated per country and for assignment accuracy results of >80% and >90% as obtained in GENECLASS2.0. The term "label" used here applies both to actual package labels and to information provided orally (as was the case of some fishmongers).

At 80% assignment accuracy, 77% of the samples with over 8 SNPs successfully amplified were assigned to the NEA and 23% of the samples with over 8 SNPs successfully amplified were assigned to the NS. Regarding the accuracy of the claimed provenance for the NEA and NS samples, 89% of the samples claimed to come from the NEA were indeed from the NEA and 33% of the samples claimed to come from the NS were from the NS (**FIGURE 5.8**). By far the most important level of mislabelling concerned samples that were claimed as coming from somewhere else (English Channel, Bay of Biscay, Galician waters, etc...) than the NEA or the NS but that did in fact come from one of these two regions. When reversing the picture and focusing solely on the samples that had been assigned to the NEA, only 59% of those samples originated from correctly labelled specimens (11% originated from specimens wrongly claimed as coming from the NS and 30% originated from specimens wrongly claimed to come from other locations). Similarly, out of the samples that had been assigned to the NS, only 20% of those samples originated from correctly labelled specimens (27% originated from specimens wrongly

claimed to come from the NEA and 53% originated from specimens wrongly claimed to come from other locations)(**FIGURE 5.8**).

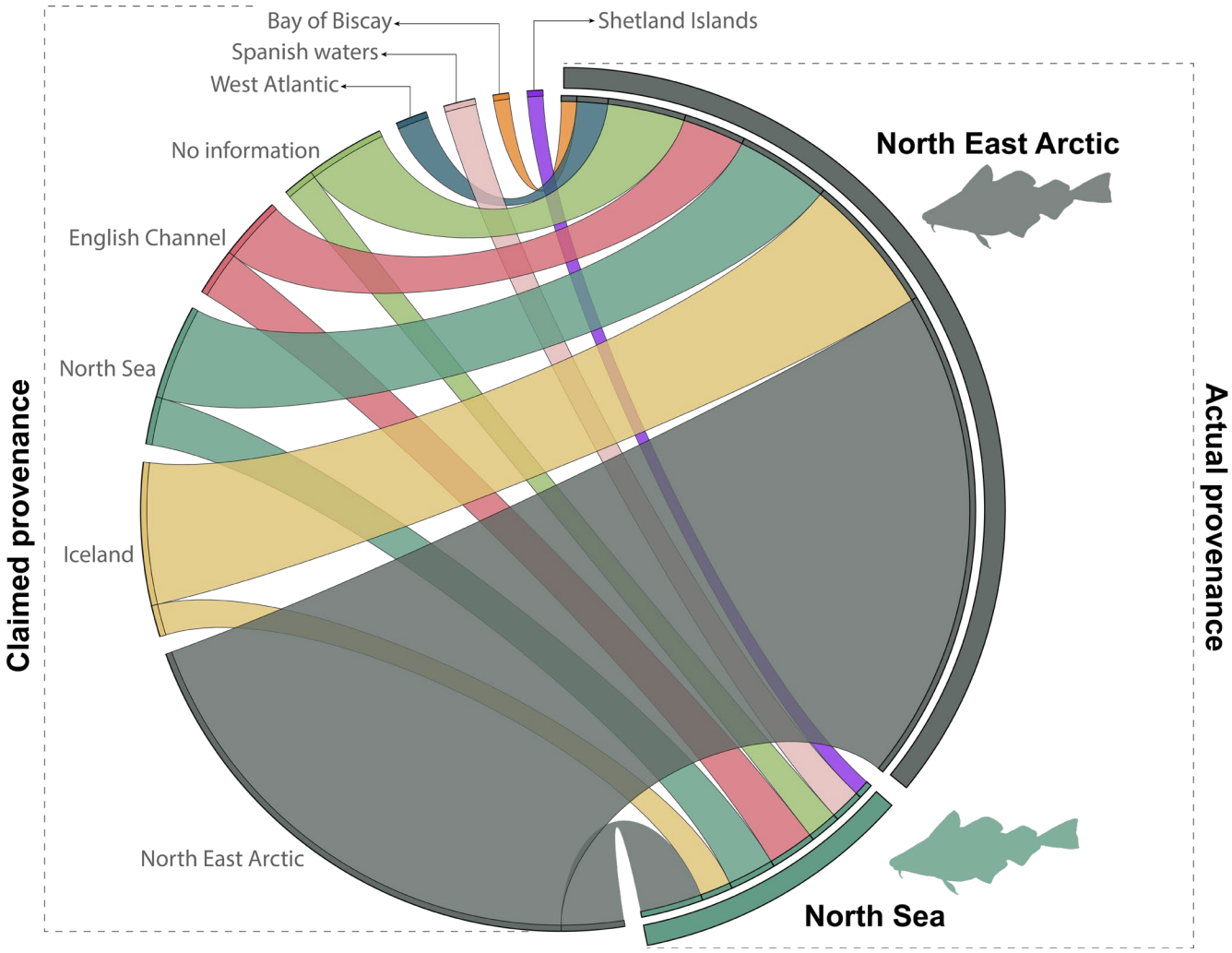


FIGURE 5.8 Circos plot illustrating the claimed provenance of cod products from retailers (verbally provided or as read on package) and the associated actual provenance as evaluated using GENECLASS2.0 with over 80% assignment accuracy.

Precision of claimed provenance had a significant effect on the observed level of mislabelling (assignment accuracy >80%) with retailers providing catch provenance with high precision (i.e. one ICES subarea) also selling the highest amount of mislabelled specimens with respect to catch area (X^2 (2, n = 70) = 18.9, p < 0.001). The level of catch provenance mislabelling was lower in products sold in supermarkets but that also

corresponded with a lower level of precision on catch provenance. A binomial logistic regression indicated that MSC certified products were significantly less mislabelled than non-MSC certified products ($p < 0.005$). When accounting for retailer type as a confounding variable (most supermarket products being MSC certified, and most fish monger products not being certified) and at >80% assignment accuracy, the MSC logo marginally affected the amount of misassignment observed ($p = 0.05$). When accounting for retailer type as a confounding variable at >90% assignment accuracy, the MSC logo no longer significantly affected the amount of misassignment observed ($p = 0.12$) (**TABLE 5.2**).

TABLE 5.2 Generalized Linear Model constructed to evaluate the relationship between correct labelling and MSC certification (Correct Label ~ MSC) with retailer type as a confounding variable (Correct Label ~ MSC + Retailer). * indicates $p \leq 0.05$ and ** indicates $p \leq 0.001$.

		Estimate	Standard Error	z value	Pr(>|z|)
Model 1 >80% assignment score - Correct Label ~ MSC	**(Intercept)**	-0.734	0.3512	-2.09	0.03662 *
	MSC - Yes	1.6278	0.5291	3.077	0.00209 **
Model 2 >80% assignment score - Correct Label ~ MSC + Retailer	**(Intercept)**	-1.2762	0.4911	-2.599	0.00936 **
	MSC - Yes	1.1664	0.5853	1.993	0.04629 *
	Retailer - SM	1.1303	0.6271	1.802	0.07148
Model 3 >90% assignment score - Correct Label ~ MSC	**(Intercept)**	-0.8267	0.4532	-1.824	0.06812
	MSC - Yes	1.8563	0.6905	2.688	0.00718 **
Model 4 >90% assignment score - Correct Label ~ MSC + Retailer	**(Intercept)**	-1.3598	0.5986	-2.272	0.0231 *
	MSC - Yes	1.2317	0.7863	1.566	0.1173
	Retailer - SM	1.3278	0.8095	1.64	0.1009

5.5 Discussion

Over the last two decades, the mislabelling of marketed seafood has raised concerns, and both academic institutions and non-governmental groups have denounced this phenomenon in an attempt to increase public awareness and provide arguments in support of better transparency in the seafood sector. In Europe, this has likely led to

improved labelling and compliance, and to reduced instances of species misdescription (Mariani et al., 2014, 2015). Similarly to other recent studies on the erroneous labelling of Atlantic cod products in Europe (Feldmann et al., 2021; Helgoe et al., 2020), we find a notably low rate of species mislabelling (~3%) of Atlantic cod. Despite this rather promising observation, scrutinizing catch location offers a much more complex and sobering picture. Overall, whether the assignment thresholds to our focal populations of origin (North East Arctic or North Sea) was 80% or 90%, around half the products appeared to be marketed under erroneous catch location information. More specifically, products sold by fish mongers were often claimed to come from local catch areas (E.g. Galician waters for cod sold in Spain, English Channel or Bay of Biscay for cod sold in France, North Sea for cod sold in Scotland) but in many cases originated from the North East Arctic (Barents Sea, Norwegian Sea, etc…) (**FIGURE 5.9**). The MSC blue ecolabel had little to no impact on the rate of mislabelling observed when accounting for the variation in mislabelling trend observed between fish mongers and supermarkets. Despite displaying slightly better trends, packaged products sold in supermarkets were still prone to high levels of catch location mislabelling. This is a serious concern for seafood transparency and traceability in the EU. If label regulations on catch provenance are to have a meaningful impact on fisheries management and consumer knowledge, simply asking to indicate ICES subareas on packages through legislation without verifying compliance might not suffice.

FIGURE 5.9 Picture from the stall of a fish monger from Quimper, France, with labels indicating that the Atlantic cod for sale originated from the subarea ICES 27.8 "Golf de Gascogne" (Bay of Biscay), this catch location was confirmed verbally by the fish monger himself. The genetic analysis indicated with over 88% assignment certainty that the cod sample collected from this fish monger was in fact from an individual caught in the North East Arctic (ICES subarea 27.1 & 27.2).

In order to coordinate coherent data collection, the FAO established fishing areas covering the world oceans. Area 27, the Northeastern part of the Atlantic Ocean comprises 14 subareas defined by ICES which are themselves divided in sub-zones (FAO, 2021). The European Commission uses these FAO areas and ICES subareas in their legislative documents as fisheries management units. Article 38 of Regulation (EU) No 1379/2013 on the common organisation of the markets in fishery and aquaculture products states that packages which contain fish including fish fillets and other fish meat, fresh, chilled or frozen that are wild caught in the Northeast Atlantic must include FAO catch area and ICES subareas on their label along with catch location information easy to interpret by consumers. The regulation was implemented in December 2014, and this is reflected in our data on labels catch information which greatly differed between products sampled in 2013 and 2020, with only 5% of the products containing ICES subarea

information on their package in 2013 as opposed to 96% in 2020. Despite providing ICES subarea information however, products differ substantially in terms of precision, with supermarket products often including several subareas spanning over several cod stocks that may display different status and catch allowance. Perhaps unsurprisingly, a lack of precision in labelling is also linked to a lower mislabelling rate highlighting a trade-off between precision of information on catch location and accuracy of product label. Notwithstanding EU incentives to increase transparency via more detailed labels, the fisheries industry might currently be ill-adapted to providing and keeping this level of information accurate throughout the supply chain, and additional steps may need to be taken to provide consumers with both precise and reliable information on the provenance of the seafood products they wish to purchase. While species identification has become common place and has permitted a large body of research to evaluate seafood species mislabelling, hardly any research has been conducted to explore the mislabelling rate of seafood catch provenance. Unlike species identification, provenance testing is still hitting many roadblocks including the development of a database extensive enough to be of practical use. More importantly, the protocols required for provenance testing are complex and require in a first instance the identification of population structures and associated stock boundaries. Due to changes in population distribution, these boundaries may change over time which adds a layer of complexity to the use of molecular tools for provenance testing. Whereas species identification using molecular tools usually generates a binary answer, the same is not true for provenance testing for which results are probabilistic in nature and certainty of assignment never reaches 100%. This, combined with lower statistical power, makes DNA-based methods less attractive when evaluating seafood catch provenance. As a result, the incentive to abide by stringent labelling regulations is lacking. In the USA, a number of court cases taking advantage of forensic DNA-based technologies for species authentication have led to prosecutions of perpetrators who were swapping seafood species for financial gain (Martinsohn et al., 2019). This along with ongoing media coverage of mislabelling occurrences largely acts as a deterrent contributing to improved species labelling compliance. On the other hand, there are few documented cases of geographical origin verification using DNA-based tools

to ensure compliance. This might partially be due to the absence of studies reporting on catch location mislabelling and to our resulting lack of understanding regarding the prevalence of this phenomenon.

The precision with which the industry reports ICES subareas should reflect the population distribution of the species being sold. Legislative bodies have used FAO areas and ICES subareas to manage their fisheries which has often raised questions over the mismatch between biological population boundaries and management areas (Reiss et al., 2009). This study uses 9 diagnostic SNPs to assign cod individuals back to either the North East Arctic or the North Sea populations. ICES, which informs the EU on management considerations, treats the North East Arctic (ICES 27.1 and 27.2) and the North Sea (ICES 27.4) cod populations as two distinct stocks, yet recognizes the existence of more than one population within these stocks. Genetic evidence supports that there are indeed two distinct groups populating these areas, but that further differentiation exists, particularly within the NEA cod stock, which is in fact composed of a migratory North East Arctic cod and a sedentary Norwegian coastal cod. These stocks are assessed separately but managed as a single unit. Similarly, the Icelandic cod stock is likely to be composed of more than one population but is managed as a single unit. In this study, we focused on management units as designated by ICES and the EU, thereby treating the NEA and the NS cod stocks as two populations without focusing on the local population units that may exist within these stocks. It should be noted however that the low assignment scores we sometimes observed to the NEA might be the product of more than one population being targetted by the fisheries industry in that management unit. A product labelled as coming from the NEA and being assigned with a low score to the NEA might therefore have originated from the sedentary Norwegian coastal cod rather than the transiant and more targetted migratory Barents sea cod. We also noted that about 34% of the products indicating that the catch location might have been the NEA also included that the product might have originated from Iceland, which is a different cod population altogether and is treated as a separate stock. This greatly blurs the picture of the catch provenance.

Samples with low assignment scores were removed from the analysis as the catch location could not be interpreted with high confidence. It is worth noting however that low assignment scores as given through GENECLASS2.0 could be the result of a number of factors and could reflect a sample caught in a population for which we did not have a reference and for which the SNP panel was not developed. We note for example that most voucher samples (17 out of 20) from the North Sea and the Barents Sea were successfully assigned back to their population of origin with over 80% accuracy. On the other hand, almost half of the samples from Svalbard (4 out of 5) displayed low assignment accuracy and one sample had to be dropped due to poor SNP amplification. Atlantic cod present in Svalbard are the result of a relatively recent climate change induced expension (Fall et al., 2018; Misund et al., 2016) and they likely originate from two ecotypes, the Northeast Arctic cod (Barents sea) and the Norwegian coastal cod (Andrade et al., 2020). We can therefore postulate that our reference samples may not have contained any Norwegian coastal cod individuals, and that low assignment might result from these cod simply originating from a co-occuring ecotype. Though we cannot say for certain that an individual with a low assignment score comes from another geographical location than the reference populations, it is likely that an individual with a low assignment score belongs to an ecotype that is not in the reference database either as a result of migration, population overlap, or lack of adequate and exhaustive sampling of the reference samples.

The majority of the cod products sampled seemed to belong to the NEA populations and whereas as lot of the cod identified as coming from the NEA had been mislabelled as coming from a different region, most of the cod that had been labelled as coming from the NEA were correctly labelled. This is in line with recent catch records reporting that 19 523 tonnes of cod from the North Sea, English Channel, and Skagerrak region were landed in 2020 (ICES, 2021a) compared to 692 903 tonnes of cod from the North East Arctic (ICES, 2021b), the region which by far provides the largest amount of Atlantic cod. A large proprotion of the alleged claims that the cod originiated from a different location than the NS or the NEA appeared generally false, which again, is in line

with the catch landings data produced by ICES. Perhaps most importantly, about 35% of the cod claimed to come from another region appeared to belong to the North Sea cod population. This deceit is particularly problematic given the current condition of the North Sea cod stock for which ICES continues to recommend reductions in catches (ICES, 2021a). The MSC having suspended their certification for all North Sea cod fisheries, it is equally important to note that 16% of the MSC certified cod labelled to come from the North East Arctic or another region seem to have in fact been caught in the North Sea (the value of 16% was true both for samples assigned with over 80% and with over 90% certainty). The high rate of cod catch location misrepresentation underscored in this study is most probably just the tip of the iceberg. Though the NEA and NS cod stocks are some of the most targetted in the Northeast Atlantic, there remain other stocks all with their unique status and quotas that could be included in such a study to improve the resolution of geographical assignment and get an exhaustive picture of the extent to which Northeast Altantic cod stocks are being targetted and mislabelled. Beyond Atlantic cod, which is a species that has been extensively studied due to its economical predominance in the fisheries industry, hundreeds of species worldwide are composed of a range of populations some differentiated due to genetic polymorphism (sympatric) others due to geographical barriers (allopatric). These populations may be managed as a single or as several stocks, may be tragetted by the fisheries industry to varying degrees, and may have drastically different conservation status. Their mismanagement or a lack of coordinated efforts to reinforce sustainable exploitation of these stocks may lead to stock collapse or extirpation (local extinction of a species). Given the rate of species mislabelling worldwide, it is equally likely that the targetted stocks are misrepresented either by mistake or to avoid stock-specific restrictions and penalties. The results depicted in this study offer a rather bleak picture of the state of seafood traceability and transparency in Europe and highlight some of the shady activities that still plague the fisheries industry and that are likely prevalent throughout the world.

The potential of using a few diagnostic SNPs to identify the population of origin of given fish samples has been recognized and demonstrated multiple times (Nielsen et al.,

2001; Nielsen, Cariani, et al., 2012; Ogden, 2008; Ogden & Linacre, 2015), including in the context of law enforcement (Martinsohn et al., 2019; Nielsen, Hemmer-Hansen, et al., 2012) and yet no market study of potential seafood point-of-origin mislabelling was ever conducted in the EU. As a result, an extensive body of research has been denouncing seafood species mislabelling over the last two decades but mislabelling of catch location continues to be essentially overlooked. We here used a typical forensic approach to assigning samples back to their population of origin by effectively testing specific alternative hypothesis with respect to the provenance of the samples (i.e. the sample comes from the NS or the samples comes from the NEA). This approach allowed us to use a relatively small number of diagnostic SNPs (9 in total) thereby greatly reducing the amount of time needed to process the samples at the expense of a more precise point-of-origin estimate which would have needed a larger panel of SNPs. Given that there remains a certain level of uncertainty with assignment scores to a given population, a greater sample size per "claim" could increase statistical power, but in the context of a market study, this is not possible. We were therefore here reliant on assignment scores leading us to drop a number of samples which did not meet the minimum threshold of 80% and 90% assignment accuracy. That being said, Nielsen et al., 2001 noted that as few as 9 diagnostic SNPs could be used reliably to tell these populations apart, and that two or three individuals could provide trustworthy and unambiguous results. Increasing sample size therefore does not necessarily imply drastically expending time and cost expenditures. Alternatively, the accuracy of assignment could potentially be improved by combining a set of methods such as genetics and stable isotope analysis (Cusa et al., 2021).

It should also be noted that, unlike biochemical tracers such as stable isotopes which vary spatially and indicate specific foraging locations, genetic based methods reflect the reproductive population of origin of a specimen (Cusa et al., 2021; Ogden & Linacre, 2015). Therefore, and despite marked genetic differences between North Sea cod and North East Arctic cod as well as strong temporal stability in their distribution (Nielsen et al., 2001), migration between populations or climate-driven distribution shifts could be

a problem when ascertaining the exact catch location of a specimen using DNA-based methods. This is rendered even more problematic when stock management boundaries do not match biological boundaries (Ogden & Linacre, 2015; Reiss et al., 2009) and calls for conservative interpretation of the results and for careful evaluation regarding the feasibility of such investigation. In the case of Atlantic cod from the NS and NEA, the main mechanism driving population structure is barriers to population interbreeding due to their diverging habitat range and preferences, migratory patterns, and reproductive locations (Hemmer-Hansen et al., 2013, 2014). Therefore, some of the loss in precision and confidence observed in our assignment probabilities most likely originates from our choice of a restricted SNP panel, which might consequently not be well suited to stand in a court of law. The genotype overlap observed in the reference population used for this analysis prevents us from drawing robust conclusions on individuals with lower assignment scores, and could also be driven by the coexistence of these two stocks at the margin of their suspected distribution, a phenomenon previously reported with other cod populations (Jorde et al., 2018). Ultimately, other molecular tools allowing for the inclusion of a more significant number of SNPs such as SNP chips might offer a much more robust insight on provenance mislabelling but they also tend to be more costly. SNP chip might be a profitable alternative if results need to be generated easily and quickly. Probe based SNP genotyping might be more practical and time saving but is not financially comparable to massively parallel sequencing which allows for the multiplexing of hundreds of specimens at relatively low cost. Each method comes with a set of advantages and disadvantages and the best tool will come down to the priorities of the user: rapidity, ease of use, accuracy, or cost.

Concerns over the environmental, economical, and human health impact of seafood mislabelling have provided momentum for a large body of studies to explore seafood markets worldwide for potential instances of misidentification and fraud. Ever improving genetic technologies and international initiatives such as the Barcode of Life have been important tools for the development of such studies, allowing for the rapid and accurate identification of fish species. On the other hand, tracing a specimen back to its population of origin with high certainty still requires substantial upstream efforts and,

consequently, forensic studies of marketed seafood point-of-origin noticeably lag behind. This study reveals an elevated level of mislabelling of cod stock provenance, which may come as a surprise given the low level of species mislabelling recently noted for Atlantic cod in Western European countries. The reasons for this high level of geographical mislabelling are unclear and could span from genuine errors due to a complex supply chain to more blatant fraudulent behaviour motivated by economic gain. Irrespective of the causes for such high mislabelling rate, it seems apparent that despite incentives for improved seafood transparency, the EU Commission must back up its legislations with verification points along the supply chain. Given the poor state of some cod stocks, authorities and retailers should be able to verify seafood provenance and customers should be given the ability to choose where their seafood comes from using reliable labels. At the moment however, catch locations on labels don't appear to offer much in terms of reliable information and consumers continue to be prey to subterfuge.

Chapter 6

General Conclusion

Traceability can be broadly defined as the ability to track back a product to its point-of-origin. This is generally maintained through the use of labels and barcodes which should enable anyone at the retailing end of the supply chain, including customers, to know precisely where their seafood originated (e.g., what species it is, where it was caught, how it was processed, what methods were used to catch it, who caught it, etc…). The amount of information provided on these labels, and therefore the degree to which a product is traceable, varies but this information is quite stringent for products caught and sold in the EU. Traceability therefore heavily relies upon the precision and accuracy of the information provided on labels, but the robustness of this process is weakened and challenged when instances of mislabelling or fraud are detected along the supply chain.

The problem of seafood mislabelling seems to have improved over time, but it is far from eliminated and will remain an issue as long as authorities struggle to verify adequately whether the industry is complying with the regulations or not. Though not all mislabelling can be attributed to pernicious incentives; other drivers, such as the lack of harmonization across regulatory bodies, poor control laboratory capacity, and the near-absence of investment in science-based methods to identify species and their point-of-origin can result in the persistence of misconduct. This thesis examined some of the societal factors that may affect seafood traceability and explored DNA-based methods that have the potential to greatly improve the continuous and regular monitoring of transparency and traceability along the supply chain. Each chapter is dedicated to a given driver which might allow mislabelling to persevere (i.e. lack of consumer knowledge, shortcomings of species identification methods, absence of a framework for the use of point-of-origin detection tools) and explores some of the associated solutions that could help strengthen the monitoring of seafood products, verify compliance, and tackle fraud in the seafood industry.

The second chapter of this thesis explores consumer's ability to identify the fish available on the market with the understanding that a lack of knowledge on species appearance and diversity is often linked to a lack of concern and empathy for these organisms. European citizens' ability to identify common commercial fish species was surprisingly low despite some cultural associations where the more commonly consumed fish were usually best recognized. Despite strict labelling requirements in the EU to improve transparency and provide consumers with comprehensive information on product characteristics, the consumers themselves seem to display little connection with the seafood they consume. Consumers are an important driver affecting market trends and associated environmental outcomes (Springmann et al., 2016; Stehfest et al., 2009), and as long as they exhibit a lack of interest for the wondrous diversity of fish being harvested and sold, the seafood industry will likely remain open to malpractice.

Despite the importance of consumer awareness, much of the responsibility for improving traceability and transparency ultimately falls on governmental agencies and on the industry itself. The third chapter therefore moves on to testing closed-tube barcoding, a promising technology for quick taxon authentication which could allow the industry and authorities to verify species identity on-site, with little expertise and laboratory capacity. Pelagic ichthyofauna was collected to demonstrate the reliability, portability and functionality of this method aboard a trawler on randomly collected species. The method proved successful in identifying a variety of commercial species in the field and even provided the possibility to identify within-species haplotypes, which may be of particular relevance for fisheries surveys. Following this validation study, the Marine Management Organization expressed great interest in implementing the method during their inspections at landing sites to prevent fraudulent substitution of species at the early nodes of the supply chain (**FIGURE 6.1**) (Graham, 2020). This chapter highlights how closed-tube barcoding could be of immense value to the seafood sector in terms of enforcement and authentication, and could easily act as a quick and reliable tool for increased seafood traceability worldwide.

FIGURE 6.1 Demonstration to the UK Marine Management Organization of quick DNA extraction and universal species identification method for the rapid, on-site authentication of fish filets (Graham, 2020).

Whereas species identification has progressed to producing quick, easy, and portable tools, forensic identification of geographical point-of-origin remains in its infancy. Notwithstanding great technological improvements allowing for the development of reliable methods to evaluate the point-of-origin of commercial products, their application is relatively complex and time consuming, and requires some level of expertise and downstream bioinformatics analysis. The fourth chapter of this thesis has evaluated some of the biological traits relevant to population genetics that could help predict whether DNA-based methods might be used as a reliable tool for the geographic validation of a given species. Through a meta-analysis exploring biological traits and genetic structure of dozens of commercial species, a framework was described to help interested parties discern whether genetic tools might be successful in verifying geographical point-of-origin

on a case-by-case basis. This study was conducted in collaboration with stable isotope experts who performed an analysis to evaluate the extent to which biochemical tracers might be used for the determination of geographical origin. It was demonstrated that despite marked progress in the development of population genetics and stable isotope tools, their application by seafood processors, retailers, government enforcement agencies, and certification bodies will require more investment and coordination to outweigh the costs involved with carrying out these types of analyses.

Traceability of the geographical origin of marketed products is a crucial element of adequate fisheries management and is increasingly sought after by investigators and management officers. When it comes to stock fishing restrictions, fraud is easy to perpetrate and difficult to detect. As demonstrated in Chapter 4 of the thesis, life history traits can be examined for any given species to evaluate if DNA-based tools are appropriate for geographical assignment. Chapter 5 of the thesis focuses on a high fecundity, demersal and somewhat migratory species that has been extensively studied, Atlantic cod. Mislabelling of seafood point-of-origin on the market remains largely unexplored and this study uses knowledge gathered in Chapter 4 as well as previously described population diagnostic markers to develop a protocol for testing the geographical origin of marketed cod and to evaluate its provenance mislabelling level. The method allowed for the successful genotyping of samples from several European countries and unveiled sobering results. The species identification method described in Chapter 3 was used to rapidly authenticate the specimens in-house and showed relatively low species mislabelling. On the other hand, mislabelling of Atlantic cod point-of-origin as indicated on packages or communicated orally by fishmongers was discouragingly high, exposing some significant loopholes in the traceability of seafood catch location. This chapter highlights the need to verify traceability claims along the supply chain and calls for studies and investigations to move beyond mere species authentication. It also demonstrates that seafood point-of-origin can be successfully evaluated in the context of traceability and that DNA-based tools can and should be used for control and enforcement activities in the fisheries industry.

Demand for seafood products is on an upward trend and the industry must continuously adapt to satisfy hungry consumers. Improved fishing technologies and increased fishing efforts have led to the collapse of once plentiful stocks on a global scale and have driven many countries to strongly regulate their fisheries. Despite these regulations however, IUU fishing activities and fraud along the supply chain persist partly because of an absence of coordinated actions to verify their successful implementation. For the last two decades, the scientific community has been using available biotechnological tools to report on the lack of seafood transparency and has noted severe challenges afflicting the industry. These types of studies continue to increase, often denouncing troubling and ongoing instances of species misrepresentation. Fewer studies have attempted to develop and list ready-to-use methods that could be directly implemented by authorities with minimal training. This thesis explored both societal and technological challenges which, if addressed with the right tools, could be solved, leading to increasing transparency and traceability. Through raising public awareness, investing in universal, quick and reliable species identification methods, and applying proposed frameworks for point-of-origin determination, governmental bodies and stakeholders have the potential to vastly improve some of the disconcerting trends currently observed in the seafood industry.

Final considerations and future directions

No one DNA-based method is perfect, and all come with a set of limitations which may render them unattractive to enforcement entities demanding for quick, simple, and universal tools. Identifying a species with high certainty still requires some level of expertise and lab work – especially as fishes are the most diverse class of vertebrates, and exploitation and trade applies to every region of the globe – but the toolkit available is increasingly robust and reliable. For example, the closed-tube barcoding method explored in Chapter 3 of this thesis and demonstrated to the Marine Management Organization in the UK (**FIGURE 6.1**) offers a relatively universal species authentication method which is quick and easy to conduct but demands some level of financial investment. This would

involve mostly the one-time construction of a comprehensive reference database of fluorescent profiles, which would quickly outweigh its financial cost through extensive practical use. The identification of species point-of-origin, a most-demanded tool for enforcement and forensic work, has proved possible with the use of DNA-based technologies on a number of occasions, leading to the development of Standard Operating Procedures by a consortium of scientists (FishPopTrace, 2013), and yet its complexity seems to have deterred agencies and stakeholders from truly exploring and implementing its applicable use. If interested parties invested in coordinated efforts to develop such methods for an increasing number of species however, this would greatly decrease the financial and time costs often associated with such methods, and the benefit of their use would quickly become apparent (Martinsohn et al., 2019).

Natural resources are under tremendous pressure and the need for good stewardship is now critical. To continue harvesting from the sea and meet rising global demands for seafood, governments around the world will need to adopt progressive and forward-looking measures, which will demand investments in environmental education of human societies, in the development of diagnostic and enforcement tools, and in the training of specialized personnel. Unsustainable and IUU fishing can lead to devastating social, economic, and ecological consequences and must be addressed urgently. This thesis explores cutting edge technologies that have the potential to transform the way the seafood industry is being monitored and regulated through their implementation. It bridges the demand for simple tools with some of the existing technologies and proposes frameworks and strategies for their adoptions in practical contexts. It is clear however that, if these tools are to be of any meaningful use, the work doesn't stop with a submitted thesis or a published peer-reviewed paper. The scientific community must continue to reach out to enforcement entities, governmental bodies, and to the industry itself to propose and demonstrate the use of these technologies and to develop coordinated frameworks for their application. This ought to be an ongoing endeavour, motivated by the prospect of a well managed sustainable industry that can continue feeding the world for generations to come.

REFERENCES

't Sas-Rolfes, M., Challender, D. W. S., Hinsley, A., Veríssimo, D., & Milner-Gulland, E. J. (2019). Illegal Wildlife Trade: Scale, Processes, and Governance. *Annual Review of Environment and Resources*, *44*, 201–228. https://doi.org/10.1146/annurev-environ-101718-033253

Abdi, H., & Valentin, D. (2007). Multiple correspondence analysis. In *Encyclopedia of measurement and statistics*. Sage.

Adobe Inc. (2019). *Adobe Illustrator*.

Agnew, D. J., Pearce, J., Pramod, G., Peatman, T., Watson, R., Beddington, J. R., & Pitcher, T. J. (2009). Estimating the worldwide extent of illegal fishing. *PLoS ONE*, *4*(2). https://doi.org/10.1371/journal.pone.0004570

Altintzoglou, T., & Nøstvold, B. H. (2014). Labelling fish products to fulfil Norwegian consumers' needs for information. *British Food Journal*, *116*(12), 1909–1920. https://doi.org/10.1108/BFJ-06-2013-0144

Anderson, J. L., Asche, F., & Garlock, T. (2018). Globalization and commoditization: The transformation of the seafood market. *Journal of Commodity Markets*, *12*(November 2017), 2–8. https://doi.org/10.1016/j.jcomm.2017.12.004

Andrade, H., van der Sleen, P., Black, B. A., Godiksen, J. A., Locke, W. L., Carroll, M. L., Ambrose, W. G., & Geffen, A. (2020). Ontogenetic movements of cod in Arctic fjords and the Barents Sea as revealed by otolith microchemistry. *Polar Biology*, *43*(5), 409–421. https://doi.org/10.1007/s00300-020-02642-1

Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]*. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Antoniou, A., & Magoulas, A. (2014). Application of Mitochondrial DNA in Stock Identi fi cation. In *Stock Identification Methods: Applications in Fishery Science* (pp. 257–295). Elsevier Inc. https://doi.org/10.1016/B978-0-12-397003-9.00013-8

Asche, F., Roll, K. H., & Trollvik, T. (2009). New aquaculture species-The whitefish market. *Aquaculture Economics and Management*, *13*(2), 76–93.

https://doi.org/10.1080/13657300902881641

Asp, E. H. (1999). *Factors affecting food decisions made by individual consumers*. *24*, 287–294.

Balmford, A., Clegg, L., Coulson, T., & Taylor, J. (2002). Why Conservationists Should Heed Pokemon. *Science*, *295*(5564), 2367–2367. https://doi.org/10.1126/science.295.5564.2367b

Barrie, J. (2018, October 15). Restaurants have been casting off catfish and whiting for cod and selling them to unwitting customers. *INews*. https://inews.co.uk/news/consumer/fish-fraud-restaurants-catfish-whiting-used-cod-seafood-208710

Barth, J. M. I., Villegas-Ríos, D., Freitas, C., Moland, E., Star, B., André, C., Knutsen, H., Bradbury, I., Dierking, J., Petereit, C., Righton, D., Metcalfe, J., Jakobsen, K. S., Olsen, E. M., & Jentoft, S. (2019). Disentangling structural genomic and behavioural barriers in a sea of connectivity. *Molecular Ecology*, *28*(6), 1394–1411. https://doi.org/10.1111/mec.15010

Bekkevold, D., Helyar, S. J., Limborg, M. T., Nielsen, E. E., Hemmer-Hansen, J., Clausen, L. A., & Carvalho, G. R. (2015). Gene-associated markers can assign origin in a weakly structured fish, Atlantic herring. *ICES Journal of Marine Science*, *72*, 1790–1801.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., & Bonhomme, F. (2004). *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II. https://kimura.univ-montp2.fr/genetix/

Bellotti, B. (2010). Food Literacy: Reconnecting the City with the Country. *Agricultural Science*, *22*(3), 29.

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D., Maes, G. E., Martinsohn, J. T., Miller, K. M., Naish, K. A., Ovenden, J. R., Primmer, C. R., Young Suk, H., Therkildsen, N. O., & Withler, R. E. (2017). Harnessing the Power of Genomics to Secure the Future of Seafood. *Trends in Ecology and Evolution*, *32*(9), 665–680. https://doi.org/10.1016/j.tree.2017.06.010

Bohonak, A. J. (1999). Dispersal, gene flow, and population structure. *The Quarterly*

Review of Biology, *74*(1), 21–45.

Bowen, G. J. (2010). Isoscapes: Spatial pattern in isotopic biogeochemistry. *Annual Review of Earth and Planetary Sciences*, *38*, 161–187. https://doi.org/10.1146/annurev-earth-040809-152429

Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*(1), 176–182. https://doi.org/10.1111/1755-0998.12428

Bradbury, I. R., Laurel, B., Snelgrove, P. V. R., Bentzen, P., & Campana, S. E. (2008). Global patterns in marine dispersal estimates: The influence of geography, taxonomic category and life history. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1644), 1803–1809. https://doi.org/10.1098/rspb.2008.0216

Bréchon, A. L., Hanner, R., & Mariani, S. (2016). A systematic analysis across North Atlantic countries unveils subtleties in cod product labelling. *Marine Policy*, *69*, 124–133. https://doi.org/10.1016/j.marpol.2016.04.014

Brunsø, K. (2003). Consumer research on fish in Europe. In *Quality of fish from catch to consumer: Labelling, monitoring and traceability*. (pp. 335–344). Wageningen Academic Publishers.

Brunsø, Karen, Verbeke, W., Olsen, S. O., & Jeppesen, L. F. (2009). Motives, barriers and quality evaluation in fish consumption situations: Exploring and comparing heavy and light users in Spain and Belgium. *British Food Journal*, *111*(7), 699–716. https://doi.org/10.1108/00070700910972387

Caballero, S., Cardeñosa, D., Soler, G., & Hyde, J. (2012). Application of multiplex PCR approaches for shark molecular identification: Feasibility and applications for fisheries management and conservation in the Eastern Tropical Pacific. *Molecular Ecology Resources*, *12*(2), 233–237. https://doi.org/10.1111/j.1755-0998.2011.03089.x

Calosso, M. C., Claydon, J. A. B., Mariani, S., & Cawthorn, D. M. (2020). Global footprint of mislabelled seafood on a small island nation. *Biological Conservation*, *245*(September 2019), 108557. https://doi.org/10.1016/j.biocon.2020.108557

Campbell, N. R., Harmon, S. A., & Narum, S. R. (2015). Genotyping-in-Thousands by

sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. *Molecular Ecology Resources*, *15*(4), 855–867. https://doi.org/10.1111/1755-0998.12357

Cardeñosa, D., Gollock, M. J., & Chapman, D. D. (2019). Development and application of a novel real-time polymerase chain reaction assay to detect illegal trade of the European eel (Anguilla anguilla) . *Conservation Science and Practice*, *1*(5), e39. https://doi.org/10.1111/csp2.39

Cardeñosa, D., Quinlan, J., Shea, K. H., & Chapman, D. D. (2018). Multiplex real-time PCR assay to detect illegal trade of CITES-listed shark species. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-34663-6

Carrera, M., & Gallardo, J. M. (2017). Determination of the Geographical Origin of All Commercial Hake Species by Stable Isotope Ratio (SIR) Analysis. *Journal of Agricultural and Food Chemistry*, *65*(5), 1070–1077. https://doi.org/10.1021/acs.jafc.6b04972

Carroll, W. F. (2009). Sushi: globalization through food culture: towards a study of global food networks. 東アジア文化交渉研究, *August*, 451–456. http://kuir.jm.kansai-u.ac.jp/dspace/handle/10112/3253

Cawthorn, D. M., Baillie, C., & Mariani, S. (2018). Generic names and mislabeling conceal high species diversity in global fisheries markets. *Conservation Letters*, *11*(5), 1–12. https://doi.org/10.1111/conl.12573

Cawthorn, D. M., & Mariani, S. (2017). Global trade statistics lack granularity to inform traceability and management of diverse and high-value fishes. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-12301-x

Chao, K.-H., Barton, K., Palmer, S., & Lanfear, R. (2020). *sangeranalyseR: simple and interactive analysis of Sanger sequencing data in R*. https://doi.org/https://doi.org/10.1101/2020.05.18.102459

Chesson, L. A., Podłesak, D. W., Thompson, A. H., Cerling, T. E., & Ehleringer, J. R. (2008). Variation of hydrogen, carbon, nitrogen, and oxygen stable isotope ratios in an american diet: Fast food meals. *Journal of Agricultural and Food Chemistry*, *56*(11), 4084–4091. https://doi.org/10.1021/jf0733618

Chesson, L. A., Valenzuela, L. O., O'Grady, S. P., Cerling, T. E., & Ehleringer, J. R. (2010). Hydrogen and Oxygen Stable Isotope Ratios of Milk in the United States. *Journal of Agricultural and Food Chemistry*, *58*(4), 2358–2363. https://doi.org/10.1021/jf904151c

Chopelet, J., Waples, R. S., & Mariani, S. (2009). Sex change and the genetic structure of marine fish populations. *Fish and Fisheries*, *10*(3), 329–343. https://doi.org/10.1111/j.1467-2979.2009.00329.x

Choquet, M., Hatlebakk, M., Dhanasiri, A. K. S., Kosobokova, K., Smolina, I., Søreide, J. E., Svensen, C., Melle, W., Kwaśniewski, S., Eiane, K., Daase, M., Tverberg, V., Skreslet, S., Bucklin, A., & Hoarau, G. (2017). Genetics redraws pelagic biogeography of Calanus. *Biology Letters*, *13*(12). https://doi.org/10.1098/rsbl.2017.0588

Christensen, R. H. B. (2019). *ordinal - Regression Models for Ordinal Data*. *R package version 2019.12-10*. https://cran.r-project.org/package=ordinal

Christiansen, H., Fournier, N., Hellemans, B., & Volckaert, F. A. M. (2018). Seafood substitution and mislabeling in Brussels' restaurants and canteens. *Food Control*, *85*, 66–75. https://doi.org/10.1016/j.foodcont.2017.09.005

Clover, C. (2013). *The End Of The Line: How Overfishing Is Changing the World and What We Eat* (2nd ed.). The New Press, New York.

Cohen, A. (1997). Sturgeon poaching and black market caviar: A case study. *Environmental Biology of Fishes*, *48*, 423–426. https://doi.org/10.1023/a:1007388803332

Cohen, N. J., Deeds, J. R., Wong, E. S., Hanner, R. H., Yancy, H. F., White, K. D., Thompson, T. M., Wahl, M., Pham, T. D., Guichard, F. M., Huh, I., Austin, C., Dizikes, G., & Gerber, S. I. (2009). Public health response to puffer fish (Tetrodotoxin) poisoning from mislabeled product. *Journal of Food Protection*, *72*(4), 810–817. https://doi.org/10.4315/0362-028X-72.4.810

Costello, C., Cao, L., Gelcich, S., Cisneros-Mata, M., Free, C. M., Froehlich, H. E., Golden, C. D., Ishimura, G., Maier, J., Macadam-Somer, I., Mangin, T., Melnychuk, M. C., Miyahara, M., de Moor, C. L., Naylor, R., Nøstbakken, L., Ojea, E., O'Reilly, E., Parma, A. M., … Lubchenco, J. (2020). The future of food from the sea. *Nature*, *June*.

https://doi.org/10.1038/s41586-020-2616-y

Cusa, M., St, K., Glew, J., Trueman, C., Buckley, L., Neat, F., Longo, C., & Mariani, S. (2021). A future for seafood point-of-origin testing using DNA and stable isotope signatures. *Reviews in Fish Biology and Fisheries*, *0123456789*. https://doi.org/10.1007/s11160-021-09680-w

De Almeida, M. D. V., Graça, P., Lappalainen, R., Giachetti, I., Kafatos, A., Remaut De Winter, A. M., & Kearney, J. M. (1997). Sources used and trusted by nationally-representative adults in the European Union for information on healthy eating. *European Journal of Clinical Nutrition*, *51*, S16–S22.

Dey, M. M., Garcia, Y. T., Kumar, P., Piumsombun, S., Haque, S., Li, L., Radam, A., & Senaratne, A. (2008). *Demand for fish in Asia : a cross-country analysis* *. 321–338. https://doi.org/10.1111/j.1467-8489.2008.00418.x

Di Muri, C., Vandamme, S. G., Peace, C., Barnes, W., & Mariani, S. (2018). Biodiversity defrosted: unveiling non-compliant fish trade in ethnic food stores. *Biological Conservation*, *217*(December 2017), 419–427. https://doi.org/10.1016/j.biocon.2017.11.028

Doherty, P. J., Planes, S., & Mather, P. (1995). Gene flow and larval duration in seven species of fish from the Great Barrier Reef. *Ecology*, *76*, 2373–2391.

Donarski, J., & Heinrich, K. (2015). *British Beef Origin Project II – Improvement of the British Beef Isotope Landscape Map (Isoscape) for Scotland and Northern Ireland*.

Donlan, C. J., & Luque, G. M. (2019). Exploring the causes of seafood fraud: A meta-analysis on mislabeling and price. *Marine Policy*, *100*, 258–264. https://doi.org/10.1016/j.marpol.2018.11.022

Doward, J. (2017, July). Think you know what fish is in your sushi? Think again. *The Guardian*. https://www.theguardian.com/lifeandstyle/2017/jul/16/sushi-fish-tropical-species-mis-sold-restaurants-shops

Drinan, D. P., Gruenthal, K. M., Canino, M. F., Lowry, D., Fisher, M. C., & Hauser, L. (2018). Population assignment and local adaptation along an isolation-by-distance gradient in Pacific cod (Gadus macrocephalus). *Evolutionary Applications*, *11*(8), 1448–1464. https://doi.org/10.1111/eva.12639

Druml, B., & Cichna-Markl, M. (2014). High resolution melting (HRM) analysis of DNA - Its role and potential in food analysis. *Food Chemistry*, *158*, 245–254. https://doi.org/10.1016/j.foodchem.2014.02.111

EUMOFA. (2017). *EU consumer habits regarding fishery and aquaculture products*. https://www.eumofa.eu/documents/20178/84590/EU+consumer+habits_final+report+.pdf/5c61348d-a69c-449e-a606-f5615a3a7e4c

EUMOFA. (2019). *The EU fish market*. https://www.eumofa.eu/documents/20178/314856/EN_The+EU+fish+market_2019.pdf/

EUMOFA. (2020). *The EU Fish Market*. https://doi.org/10.2771/664425

Eurostat. (2020). *Urban and rural living in the EU*. https://ec.europa.eu/eurostat/web/products-eurostat-news/-/EDN-20200207-1

Fall, J., Ciannelli, L., Skaret, G., & Johannesen, E. (2018). Seasonal dynamics of spatial distributions and overlap between Northeast Arctic cod (Gadus morhua) and capelin (Mallotus villosus) in the Barents Sea. *PLoS ONE*, *13*(10), 1–26. https://doi.org/10.1371/journal.pone.0205921

U.N. FAO. Codex general standard for the labelling of prepackaged foods, in CODEX STAN 1985, (1985) (testimony of FAO).

FAO. (2017). *Fisheries and aquaculture software. FishStat Plus - Universal software for fishery statistical time series.* FAO Fisheries and Aquaculture Department [Online]. Rome. http://www.fao.org/fishery/statistics/software/fishstat/en

FAO. (2018a). Overview of food fraud in the fisheries sector. In *Document prepared by Alan Reilly* (Vol. 1165, Issue 1165). http://www.fao.org/3/i8791en/I8791EN.pdf

FAO. (2018b). The State Of World Fisheries And Aquaculture 2018 - Meeting the sustainable goals. *Rome. Licence: CC BY-NC-SA 3.0 IGO*, *29*. https://doi.org/10.1111/fog.12466

FAO. (2020). *The State of World Fisheries and Aquaculture 2020*. Sustainability in Action. Rome.

FAO. (2021). *FAO Major Fishing Areas. ATLANTIC, NORTHEAST (Major Fishing Area 27). CWP Data Collection.FAO Major Fishing Areas. ATLANTIC, NORTHEAST (Major Fishing*

Area 27). CWP Data Collection. In: FAO Fisheries Division [Online]. Rome. http://www.fao.org/fishery/area/Area27/en

FDA. (2020). *Guidance for Industry: the Seafood List. U.S. Food and Drug Administration*. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-industry-seafood-list

Feldman, K. A., Werner, S. B., Cronan, S., Hernandez, M., Horvath, A. R., Lea, C. S., Au, A. M., & Vugia, D. J. (2005). A large outbreak of scombroid fish poisoning associated with eating escolar fish (Lepidocybium flavobrunneum). *Epidemiology and Infection*, *133*(1), 29–33. https://doi.org/10.1017/S095026880400322X

Feldmann, F., Ardura, A., Blanco-Fernandez, C., & Garcia-Vazquez, E. (2021). Dna analysis detects different mislabeling trend by country in european cod fillets. *Foods*, *10*(7). https://doi.org/10.3390/foods10071515

FERA. (2020). https://www.fera.co.uk/food-safety/authenticity-testing

Fernandes, T. J. R., Amaral, J. S., & Mafra, I. (2020). DNA barcode markers applied to seafood authentication: an updated review. *Critical Reviews in Food Science and Nutrition*, *0*(0), 1–32. https://doi.org/10.1080/10408398.2020.1811200

Fernandes, T. J. R., Silva, C. R., Costa, J., Oliveira, M. B. P. P., & Mafra, I. (2017). High resolution melting analysis of a COI mini-barcode as a new approach for Penaeidae shrimp species discrimination. *Food Control*, *82*, 8–17. https://doi.org/10.1016/j.foodcont.2017.06.016

Fields, A. T., Fischer, G. A., Shea, S. K. H., Zhang, H., Feldheim, K. A., & Chapman, D. D. (2020). DNA Zip-coding: identifying the source populations supplying the international trade of a critically endangered coastal shark. *Animal Conservation*, 1–9. https://doi.org/10.1111/acv.12585

FishPopTrace. (2013). *Traceability of Fish Populations and Fish Products - Advances and Contribution to Sustainable Fisheries*.

Fox, M., Mitchell, M., Dean, M., Elliott, C., & Campbell, K. (2018). The seafood supply chain from a fraudulent perspective. *Turkish Bulletin of Hygiene and Experimental Biology*, *75*(10), 939–963. https://doi.org/10.5505/turkhijyen.2018.75508

Freeland, J., Kirk, H., & Petersen, S. (2005). Molecular markers in ecology. In *Molecular*

Ecology (pp. 31–62).

Freeland, J. R., & Petersen, S. D. (2011). *Molecular Ecology John* (2nd ed.). Wiley & Sons.

Galarza, J. A., Carreras-Carbonell, J., Macpherson, E., Pascual, M., Roques, S., Turner, G. F., & Rico, C. (2009). The influence of oceanographic fronts and early-life-history traits on connectivity among littoral fish species. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(5), 1473–1478. https://doi.org/10.1073/pnas.0806804106

Gallic, B. Le, & Cox, A. (2006). An economic analysis of illegal, unreported and unregulated (IUU) fishing: Key drivers and possible solutions. *Marine Policy*, *30*(6), 689–695. https://doi.org/10.1016/j.marpol.2005.09.008

Gaviglio, A., Demartini, E., Mauracher, C., & Pirani, A. (2014). Consumer perception of different species and presentation forms of fish: An empirical analysis in Italy. *Food Quality and Preference*, *36*, 33–49. https://doi.org/10.1016/j.foodqual.2014.03.002

Gilbey, J., Coughlan, J., Wennevik, V., Prodöhl, P., Stevens, J. R., Garcia De Leaniz, C., Ensing, D., Cauwelier, E., Cherbonnel, C., Consuegra, S., Coulson, M. W., Cross, T. F., Crozier, W., Dillane, E., Ellis, J. S., García-Vázquez, E., Griffiths, A. M., Gudjonsson, S., Hindar, K., … Verspoor, E. (2018). A microsatellite baseline for genetic stock identification of European Atlantic salmon (Salmo salar L.). *ICES Journal of Marine Science*, *75*(2), 662–674. https://doi.org/10.1093/icesjms/fsx184

Glover, K. A., Quintela, M., Wennevik, V., Besnier, F., Sørvik, A. G. E., & Skaala, Ø. (2012). Three decades of farmed escapees in the wild: A spatio-temporal analysis of atlantic salmon population genetic structure throughout norway. *PLoS ONE*, *7*(8). https://doi.org/10.1371/journal.pone.0043129

Gong, Y., Li, Y., Chen, X., & Chen, L. (2018). Potential use of stable isotope and fatty acid analyses for traceability of geographic origins of jumbo squid (Dosidicus gigas). *Rapid Communications in Mass Spectrometry*, *32*(7), 583–589. https://doi.org/10.1002/rcm.8071

Gopi, K., Mazumder, D., Sammut, J., & Saintilan, N. (2019). Determining the provenance and authenticity of seafood: a review of current methodologies. *Trends in Food Science & Technology*.

Gopi, Karthik, Mazumder, D., Sammut, J., & Saintilan, N. (2019). Determining the provenance and authenticity of seafood: A review of current methodologies. *Trends in Food Science and Technology*, *91*(May), 294–304. https://doi.org/10.1016/j.tifs.2019.07.010

Gordoa, A., Carreras, G., Sanz, N., & Viñas, J. (2017). Tuna species substitution in the Spanish commercial chain: A knock-on effect. *PLoS ONE*, *12*(1), 1–15. https://doi.org/10.1371/journal.pone.0170809

Gouda, S., Kerry, R. G., Das, A., & Chauhan, N. S. (2020). Wildlife forensics: A boon for species identification and conservation implications. *Forensic Science International*, *317*, 110530. https://doi.org/10.1016/j.forsciint.2020.110530

Graham, C. (2020). *DNA testing to establish fish species*. Marine Management Organisation. https://marinedevelopments.blog.gov.uk/2020/02/14/dna-testing-to-establish-fish-species/

Griffiths, A. M., Miller, D. D., Egan, A., Fox, J., Greenfield, A., & Mariani, S. (2013). DNA barcoding unveils skate (Chondrichthyes: Rajidae) species diversity in 'ray' products sold across Ireland and the UK. *PeerJ*, *1*, e129. https://doi.org/10.7717/peerj.129

Handy, S. M., Mueller, S., Jacob, S. M., Paul, S. Z., Garrett, S. D., & Deeds, J. R. (2017). Evaluation of the Agilent Technologies Bioanalyzer-based DNA Fish Identification Solution. *Food Control*, *73*, 627–633. https://doi.org/10.1016/j.foodcont.2016.09.013

Hanner, R., Becker, S., Ivanova, N. V., & Steinke, D. (2011). FISH-BOL and seafood identification: Geographically dispersed case studies reveal systemic market substitution across Canada. *Mitochondrial DNA*, *22*(SUPPL. 1), 106–122. https://doi.org/10.3109/19401736.2011.588217

Harvey, A. C., Glover, K. A., Taylor, M. I., Creer, S., & Carvalho, G. R. (2015). A common garden design reveals population-specific variability in potential impacts of hybridization between populations of farmed and wild Atlantic salmon, Salmo salar L. *Evolutionary Applications*, *9*(3), 435–449. https://doi.org/10.1111/eva.12346

Heaton, M. P., Keen, J. E., Clawson, M. L., Harhay, G. P., Bauer, N., Shultz, C., Green, B. T., Durso, L., Chitko-McKown, C. G., & Laegreid, W. W. (2005). Use of bovine single nucleotide polymorphism markers to verify sample tracking in beef processing. *Public*

Veterinary Medicine, *226*(8), 1311–1314.

Heaton, M. P., Leymaster, K. A., Kalbfleisch, T. S., Kijas, J. W., Clarke, S. M., McEwan, J., Maddox, J. F., Basnayake, V., Petrik, D. T., Simpson, B., Smith, T. P. L., & Chitko-Mckown, C. G. (2014). SNPs for parentage testing and traceability in globally diverse breeds of sheep. *PLoS ONE*, *9*(4), 1–10. https://doi.org/10.1371/journal.pone.0094851

Hebert, P. D. N., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, *270*, 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert, P. D. N., Ratnasingham, S., & DeWaard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences*, *270*(SUPPL. 1), 1–4. https://doi.org/10.1098/rsbl.2003.0025

Helgoe, J., Oswald, K. J., & Quattro, J. M. (2020). A comprehensive analysis of the mislabeling of Atlantic cod (Gadus morhua) products in Spain. *Fisheries Research*, *222*(November 2019), 105400. https://doi.org/10.1016/j.fishres.2019.105400

Hemmer-Hansen, J., Nielsen, E. E., Therkildsen, N. O., Taylor, M. I., Ogden, R., Geffen, A. J., Bekkevold, D., Helyar, S., Pampoulie, C., Johansen, T., & Carvalho, G. R. (2013). A genomic island linked to ecotype divergence in Atlantic cod. *Molecular Ecology*, *22*(10), 2653–2667. https://doi.org/10.1111/mec.12284

Hemmer-Hansen, J., Therkildsen, N. O., Meldrup, D., & Nielsen, E. E. (2014). Conserving marine biodiversity: Insights from life-history trait candidate genes in Atlantic cod (Gadus morhua). *Conservation Genetics*, *15*(1), 213–228. https://doi.org/10.1007/s10592-013-0532-5

Hicks, C. C., Cohen, P. J., Graham, N. A. J., Nash, K. L., Allison, E. H., D'Lima, C., Mills, D. J., Roscher, M., Thilsted, S. H., Thorne-Lyman, A. L., & MacNeil, M. A. (2019). Harnessing global fisheries to tackle micronutrient deficiencies. *Nature*, *574*(7776), 95–98. https://doi.org/10.1038/s41586-019-1592-6

Hilborn, R., Amoroso, R. O., Anderson, C. M., Baum, J. K., Branch, T. A., Costello, C., De Moor, C. L., Faraj, A., Hively, D., Jensen, O. P., Kurota, H., Little, L. R., Mace, P.,

McClanahan, T., Melnychuk, M. C., Minto, C., Osio, G. C., Parma, A. M., Pons, M., … Ye, Y. (2020). Effective fisheries management instrumental in improving fish stock status. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(4), 2218–2224. https://doi.org/10.1073/pnas.1909726116

Hilborn, R., & Hilborn, U. (2019). *Ocean Recovery*. Oxford University Press.

Hill, J., Demarest, B., Bisgrove, B., Su, Y., Smith, M., & Yost, H. (2014). Poly peak parser: Method and software for identification of unknown indels using sanger sequencing of polymerase chain reaction products. *Developmental Dynamics*. https://doi.org/10.1002/dvdy.24183

Ho, J. K. I., Puniamoorthy, J., Srivathsan, A., & Meier, R. (2020). MinION sequencing of seafood in Singapore reveals creatively labelled flatfishes, confused roe, pig DNA in squid balls, and phantom crustaceans. *Food Control*, *112*(November 2019), 107144. https://doi.org/10.1016/j.foodcont.2020.107144

Hooykaas, M. J. D., Schilthuizen, M., Aten, C., Hemelaar, E. M., Albers, C. J., & Smeets, I. (2019). Identification skills in biodiversity professionals and laypeople: A gap in species literacy. *Biological Conservation*, *238*(July), 108202. https://doi.org/10.1016/j.biocon.2019.108202

Horreo, J. L., Machado-Schiaffino, G., Ayllon, F., Griffiths, A. M., Bright, D., Stevens, J. R., & Garcia-Vazquez, E. (2011). Impact of climate change and human-mediated introgression on southern European Atlantic salmon populations. *Global Change Biology*, *17*(5), 1778–1787. https://doi.org/10.1111/j.1365-2486.2010.02350.x

Hutchinson, W. F., Carvalho, G. R., & Rogers, S. I. (2001). Marked genetic structuring in localised spawning populations of cod Gadus morhua in the North Sea and adjoining waters, as revealed by microsatellites. *Marine Ecology - Progess Series*, *223*, 251–260.

ICES. (2019). ICES Annual Report. *ICES Annual Report: Copenhagen*. https://doi.org/10.17895/ices.pub.5967

ICES. (2021a). *Cod (Gadus morhua) in Subarea 4, Division 7.d, and Subdivision 20 (North Sea, eastern English Channel, Skagerrak)*.

ICES. (2021b). *Cod (Gadus morhua) in subareas 1 and 2 (Northeast Arctic)*.

https://doi.org/, https://doi.org/10.17895/ices.advice.7741

Ivanova, N. V., Zemlak, T. S., Hanner, R. H., & Hebert, P. D. N. (2007). Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes*, *7*(4), 544–548. https://doi.org/10.1111/j.1471-8286.2007.01748.x

Jacquet, J. L., & Pauly, D. (2008). Trade secrets: Renaming and mislabeling of seafood. *Marine Policy*, *32*(3), 309–318. https://doi.org/10.1016/j.marpol.2007.06.007

Johansen, T., Besnier, F., Quintela, M., Jorde, P. E., Glover, K. A., Westgaard, J. I., Dahle, G., Lien, S., & Kent, M. P. (2020). Genomic analysis reveals neutral and adaptive patterns that challenge the current management regime for East Atlantic cod Gadus morhua L. *Evolutionary Applications*, *13*(10), 2673–2688. https://doi.org/10.1111/eva.13070

Johri, S., Solanki, J., Cantu, V. A., Fellows, S. R., Edwards, R. A., Moreno, I., Vyas, A., & Dinsdale, E. A. (2019). 'Genome skimming' with the MinION hand-held sequencer identifies CITES-listed shark species in India's exports market. *Scientific Reports*, *9*(1), 1–13. https://doi.org/10.1038/s41598-019-40940-9

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jones, M. (2004). *Cultured Aquatic Species Information Programme. Salmo salar. Cultured Aquatic Species Information Programme*. FAO Fisheries Division [Online].

Jónsdóttir, Ó. D. B., Imsland, A. K., Atladóttir, Ó. Y., & Daníelsdóttir, A. K. (2003). Nuclear DNA RFLP variation of Atlantic cod in the North Atlantic Ocean. *Fisheries Research*, *63*(3), 429–436. https://doi.org/10.1016/S0165-7836(03)00098-5

Jorde, P. E., Kleiven, A. R., Sodeland, M., Olsen, E. M., Ferter, K., Jentoft, S., Knutsen, H., & Grant, W. S. (2018). Who is fishing on what stock: Population-of-origin of individual cod (Gadus morhua) in commercial and recreational fisheries. *ICES Journal of Marine Science*, *75*(6), 2153–2162. https://doi.org/10.1093/icesjms/fsy080

Kallet, A., & Schlink, F. J. (1933). *100,000,000 Guinea Pigs: Dangers in Everyday Foods, Drugs, and Cosmetics*. Grosset & Dunlap.

Kassambara, A. (2020). *ggpubr: "ggplot2" Based Publication Ready Plots. R package version 0.2.5*. https://cran.r-project.org/package=ggpubr

Kassambara, A., & Mundt, F. (2017). Package 'factoextra'. *Extract and Visualize the Results of Multivariate Data Analyses*, *76*.

Kehagia, O., Chrysochou, P., Chryssochoidis, G., Krystallis, A., & Linardakis, M. (2007). European consumers' perceptions, definitions and expectations of traceability and the importance of labels, and the differences in these perceptions by product type. *Sociologia Ruralis*, *47*(4), 400–416. https://doi.org/10.1111/j.1467-9523.2007.00445.x

Kelly, S., Heaton, K., & Hoogewerff, J. (2005). Tracing the geographical origin of food: The application of multi-element and multi-isotope analysis. *Trends in Food Science & Technology*, *16*, 155–167.

Kim, H., Suresh Kumar, K., & Shin, K. H. (2015). Applicability of stable C and N isotope analysis in inferring the geographical origin and authentication of commercial fish (Mackerel, Yellow Croaker and Pollock). *Food Chemistry*, *172*, 523–527. https://doi.org/10.1016/j.foodchem.2014.09.058

Kinlan, B. P., & Gaines, S. D. (2003). Propagule dispersal in marine and terrestrial environments: A community perspective. *Ecology*, *84*(8), 2007–2020. https://doi.org/10.1890/01-0622

Kroetz, K., Luque, G. M., Gephart, J. A., Jardine, S. L., Lee, P., Chicojay Moore, K., Cole, C., Steinkruger, A., & Donlan, C. J. (2020). Consequences of seafood mislabeling for marine populations and fisheries management. *Proceedings of the National Academy of Sciences*, *117*(48), 30318–30323. https://doi.org/10.1073/pnas.2003741117

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J., & Marra, M. A. (2009). Circos: An information aesthetic for comparative genomics. *Genome Research*, *19*(9), 1639–1645. https://doi.org/10.1101/gr.092759.109

Kurlansky, M. (1997). *Cod: A Biography of the Fish that Changed the World*. Vintage/Ebury.

Labelfish Consortium. (2014). *Standard Operating Procedure for the Genetic Identification of Fish Species Using Dna Barcoding (Mitochondrial Cytochrome-C-Oxidase I Sequencing)* (Issue December).

Lam, V. (2007). Oilfish - the case of the imitation Atlantic cod. *Sea Around Us Project Newsletter*, *40*. http://www.seaaroundus.org/newsletter/Issue40.pdf

Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, *30*(22), 3276–3278.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software*, *25*, 1–18.

Leahy, S. (2021). Revealed: seafood fraud happening on a vast global scale. *The Guardian*. https://www.theguardian.com/environment/2021/mar/15/revealed-seafood-happening-on-a-vast-global-scale

Leal, M. C., Pimentel, T., Ricardo, F., Rosa, R., & Calado, R. (2015). Seafood traceability: Current needs, available tools, and biotechnological challenges for origin certification. *Trends in Biotechnology*, *33*(6), 331–336. https://doi.org/10.1016/j.tibtech.2015.03.003

Levin, P. S., & Dufault, A. (2010). Eating up the food web. *Fish and Fisheries*, *11*(3), 307–312. https://doi.org/10.1111/j.1467-2979.2010.00355.x

Luque, G. M., & Donlan, C. J. (2019). The characterization of seafood mislabeling: A global meta-analysis. *Biological Conservation*, *236*(January), 556–570. https://doi.org/10.1016/j.biocon.2019.04.006

Magoč, T., & Salzberg, S. L. (2011). FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, *27*(21), 2957–2963. https://doi.org/10.1093/bioinformatics/btr507

Magoulas, A., Castilho, R., Caetano, S., Marcato, S., & Patarnello, T. (2006). Mitochondrial DNA reveals a mosaic pattern of phylogeographical structure in Atlantic and Mediterranean populations of anchovy (Engraulis encrasicolus). *Molecular Phylogenetics and Evolution*, *39*(3), 734–746. https://doi.org/10.1016/j.ympev.2006.01.016

Mariani, S., & Bekkevold, D. (2014). The Nuclear Genome Neutral and Adaptive Markers in Fisheries Science. In Elsevier (Ed.), *Stock Identification Methods: Applications in Fishery Science* (pp. 297–327). Academic Press. https://doi.org/https://doi.org/10.1016/C2011-0-07625-1

Mariani, S., Ellis, J., O'Reilly, A., Bréchon, A. L., Sacchi, C., & Miller, D. D. (2014). Mass media influence and the regulation of illegal practices in the seafood market.

Conservation Letters, *7*(5), 478–483. https://doi.org/10.1111/conl.12085

Mariani, S., Griffiths, A. M., Velasco, A., Kappel, K., Jerome, M., Perez-Martin, R. I., Schroder, U., Verrez-Bagnis, V., Silva, H., Vandamme, S. G., Boufana, B., Mendes, R., Shorten, M., Smith, C., Hankard, E., Hook, S. A., Weymer, A. S., Gunning, D., & Sotelo, C. G. (2015). Low mislabeling rates indicate marked improvements in European seafood market operations. *Frontiers in Ecology and the Environment*, *13*(10), 536–540. https://doi.org/10.1890/150119

Marko, P. B., Nance, H. A., & Guynn, K. D. (2011). Genetic detection of mislabeled fish from a certified sustainable fishery. *Current Biology*, *21*(16), R621–R622. https://doi.org/10.1016/j.cub.2011.07.006

Martinsohn, J. T., Raymond, P., Knott, T., Glover, K. A., Nielsen, E. E., Eriksen, L. B., Ogden, R., Casey, J., & Guillen, J. (2019). DNA-analysis to monitor fisheries and aquaculture: Too costly? *Fish and Fisheries*, *20*(2), 391–401. https://doi.org/10.1111/faf.12343

McMahon, K. W., Hamady, L. L., & Thorrold, S. R. (2013). A review of ecogeochemistry approaches to estimating movements of marine animals. *Limnology and Oceanography*, *58*(2), 697–714. https://doi.org/10.4319/lo.2013.58.2.0697

Menegon, M., Cantaloni, C., Rodriguez-Prieto, A., Centomo, C., Abdelfattah, A., Rossato, M., Bernardi, M., Xumerle, L., Loader, S., & Delledonne, M. (2017). On site DNA barcoding by nanopore sequencing. *PLoS ONE*, *12*(10), 1–18. https://doi.org/10.1371/journal.pone.0184741

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *5*(6). https://doi.org/10.1101/pdb.prot5448

Miller, D. D., Clarke, M., & Mariani, S. (2012). Mismatch between fish landings and market trends: A western European case study. *Fisheries Research*, *121–122*, 104–114. https://doi.org/10.1016/j.fishres.2012.01.016

Miller, D. D., Jessel, A., & Mariani, S. (2012). Seafood mislabelling: Comparisons of two western European case studies assist in defining influencing factors, mechanisms and motives. *Fish and Fisheries*, *13*(3), 345–358. https://doi.org/10.1111/j.1467-2979.2011.00426.x

Miller, D. D., & Mariani, S. (2010). Smoke, mirrors, and mislabeled cod: Poor transparency in the European seafood industry. *Frontiers in Ecology and the Environment*, *8*(10), 517–521. https://doi.org/10.1890/090212

Miller, D. D., & Sumaila, U. R. (2016). IUU Fishing and Impact on the Seafood Industry. In *Seafood Authenticity and Traceability* (pp. 83–95). Elsevier Inc. https://doi.org/10.1016/B978-0-12-801592-6.00004-8

Misund, O. A., Heggland, K., Skogseth, R., Falck, E., Gjøsæter, H., Sundet, J., Watne, J., & Lønne, O. J. (2016). Norwegian fisheries in the Svalbard zone since 1980. Regulations, profitability and warming waters affect landings. *Polar Science*, *10*(3), 312–322. https://doi.org/10.1016/j.polar.2016.02.001

Mohneke, M., Erguvan, F., & Schlüter, K. (2016). Explorative study about knowledge of species in the field of early years education. *J. Emergent Sci.*, *11*, 1–234.

Morin, J. F., & Lees, M. (2018). *Food Integrity Handbook. A guide to food authenticity issues and analytical solutions*.

MSC. (2019). *North Sea cod to lose sustainability certification*. https://www.msc.org/media-centre/press-releases/press-release/north-sea-cod-to-lose-sustainability-certification

MSC. (2020). *MSC 2019–20 Annual Report Supplementary Information*. https://www.msc.org/docs/default-source/default-document-library/about-the-msc/msc-2019-20-annual-report-supplementary-information.pdf?sfvrsn=68e38364_6

Murray, A. D., & Fofana, A. (2002). The changing nature of UK fish retailing. *Marine Resource Economics*, *17*, 335–339.

Murray, R. (2009). *The end of the line*. https://theendofthelinemovie.com/

Naaum, A. M., Cusa, M., Singh, M., Bleicher, Z., Elliott, C., Goodhead, I. B., Hanner, R. H., Helyar, S. J., Mariani, S., Rice, J. E., Wangh, L. J., & Sanchez, J. A. (2021). Validation of FASTFISH-ID: A new commercial platform for rapid fish species authentication via universal closed-tube barcoding. *Food Research International*, *141*, 110035. https://doi.org/10.1016/j.foodres.2020.110035

Naaum, A. M., Warner, K., Mariani, S., Hanner, R. H., & Carolin, C. D. (2016). Seafood

Mislabeling Incidence and Impacts. In *Seafood Authenticity and Traceability: A DNA-based Pespective*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-801592-6.00001-2

Nestle, M., Co-chair, N. Y., Wing, R., Disogra, L., Foods, D., Ca, S. M., Drewnowski, A., Mi, A. A., Middleton, S., Oh, C., Sigman-grant, M., Sobal, J., Winston, M., Association, A. H., Tx, D., & Economos, C. (1998). Behavioral and Social Influences on Food Choice. *Nutrition Reviews*, *56*(5), S50–S64.

Nielsen, E. E. (2016). Chapter 8 - Population or Point-of-Origin Identification. In *Seafood Authenticity and Traceability: A DNA-based Pespective*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-801592-6.00008-5

Nielsen, E. E., Cariani, A., Aoidh, E. M., Maes, G., Milano, I., Ogden, R., Taylor, M., Hemmer-Hansen, J., Babbucci, M., Bargelloni, L., Bekkevold, D., Diopere, E., Grenfell, L., Helyar, S., Limborg, M. T., Martinsohn, J. T., McEwing, R., Panitz, F., Patarnello, T., … Carvalho, G. R. (2012). Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nature Communications*, *3*, 851–856. https://doi.org/10.1038/ncomms1845

Nielsen, E. E., Hansen, M. M., Schmidt, C., Meldrup, D., & Grønkjaer, P. (2001). Population of origin of atlantic cod. *Nature*, *413*(6853), 272. https://doi.org/10.1038/35095112

Nielsen, E. E., Hemmer-Hansen, J., & Bekkevold, D. (2012). Development and application of molecular tools to investigate the mislabeling of cod sold in Sweden. In *Case Studies in Food Safety and Authenticity* (pp. 326–333). Woodhead Publishing Limited. https://doi.org/10.1533/9780857096937.6.326

Nielsen, E. E., Hemmer-Hansen, J., Poulsen, N. A., Loeschcke, V., Moen, T., Johansen, T., Mittelholzer, C., Taranger, G. L., Ogden, R., & Carvalho, G. R. (2009). Genomic signatures of local directional selection in a high gene flow marine organism; The Atlantic cod (Gadus morhua). *BMC Evolutionary Biology*, *9*(1), 1–11. https://doi.org/10.1186/1471-2148-9-276

O'Connor, S., Ono, R., & Clarkson, C. (2011). Pelagic fishing at 42,000 years before the present and the maritime skills of modern humans. *Science*, *334*(November), 1117–1121.

O'Leary, D. B., Coughlan, J., Dillane, E., McCarthy, T. V., & Cross, T. F. (2007). Microsatellite variation in cod Gadus morhua throughout its geographic range. *Journal of Fish Biology*, *70*(SUPPL. C), 310–335. https://doi.org/10.1111/j.1095-8649.2007.01451.x

OCEANA. (2016). Deceptive Dishes : Seafood Swaps Found Worldwide Table of Contents. In *Report prepared by Warner, Kimberly; Lowell, Beth; Geren, Sarah; and Talmage, Spencer*.

OECD. (2018). *Combatting illegal, unreported and unregulated fishing. Where countries stand and where efforts should concentrate in the future*. https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=TAD/FI(2017)16/FINAL&docLanguage=En

Ogden, R. (2008). Fisheries forensics: The use of DNA tools for improving compliance, traceability and enforcement in the fishing industry. *Fish and Fisheries*, *9*(4), 462–472. https://doi.org/10.1111/j.1467-2979.2008.00305.x

Ogden, R., & Linacre, A. (2015). Wildlife forensic science: A review of genetic geographic origin assignment. *Forensic Science International: Genetics*, *18*(September 2009), 152–159. https://doi.org/10.1016/j.fsigen.2015.02.008

Ogden, R., & Murray-Dickson, G. (2014). DEFRA - REPORT - Geographic traceability tools for commercial fish and fish products. In *Defra Fisheries Traceability Project*.

Ogle, D. H. P., Wheeler, P., & Dinno, A. (2020). *FSA: Fisheries Stock Analysis. R package version 0.8.30*. https://github.com/droglenc/FSA

Olson, J., Clay, P. M., & Pinto da Silva, P. (2014). Putting the seafood in sustainable food systems. *Marine Policy*, *43*, 104–111. https://doi.org/10.1016/j.marpol.2013.05.001

Palsbøll, P. J., Bérubé, M., & Allendorf, F. W. (2007). Identification of management units using population genetic data. *Trends in Ecology and Evolution*, *22*(1), 11–16. https://doi.org/10.1016/j.tree.2006.09.003

Palumbi, S. R. (2003). Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications*, *13*, 146–158.

Pappalardo, A. M., Federico, C., Sabella, G., Saccone, S., & Ferrito, V. (2015). A COI nonsynonymous mutation as diagnostic tool for intraspecific discrimination in the

European anchovy Engraulis encrasicolus (linnaeus). *PLoS ONE*, *10*(11), 1–12. https://doi.org/10.1371/journal.pone.0143297

Park, J., Lee, J., Seto, K., Hochberg, T., Wong, B. A., Miller, N. A., Takasaki, K., Kubota, H., Oozeki, Y., Doshi, S., Midzik, M., Hanich, Q., Sullivan, B., Woods, P., & Kroodsma, D. A. (2020). Illuminating dark fishing fleets in North Korea. *Science Advances*, *6*(30), 1–8. https://doi.org/10.1126/sciadv.abb1197

Paton, T. F., Marr, I., O'Keefe, Z., & Inglis, T. J. J. (2021). Development, deployment and in-field demonstration of mobile coronavirus SARS-CoV-2 Nucleic acid amplification test. *Journal of Medical Microbiology*, *70*(4). https://doi.org/10.1099/JMM.0.001346

Pauly, D. (2019). *Vanishing Fish: Shifting Baselines and the Future of Global Fisheries*. Greystone Books.

Perry, E. A., Thomas, H., Samra, H. R., Edmonstone, S., Davidson, L., Faulkner, A., Petermann, L., Manafò, E., & Kirkpatrick, S. I. (2017). Identifying attributes of food literacy: A scoping review. *Public Health Nutrition*, *20*(13), 2406–2415. https://doi.org/10.1017/S1368980017001276

Petrou, E. L., Drinan, D. P., Kopperl, R., Lepofsky, D., Yang, D., Moss, M. L., & Hauser, L. (2019). Intraspecific DNA contamination distorts subtle population structure in a marine fish: Decontamination of herring samples before restriction-site associated sequencing and its effects on population genetic statistics. *Molecular Ecology Resources*, *19*(5), 1131–1143. https://doi.org/10.1111/1755-0998.12978

Pieniak, Z., Vanhonacker, F., & Verbeke, W. (2013). Consumer knowledge and use of information about fish and aquaculture. *Food Policy*, *40*, 25–30. https://doi.org/10.1016/j.foodpol.2013.01.005

Pieniak, Z., Verbeke, W., Scholderer, J., Brunsø, K., & Olsen, S. O. (2007). European consumers' use of and trust in information sources about fish. *Food Quality and Preference*, *18*(8), 1050–1063. https://doi.org/10.1016/j.foodqual.2007.05.001

Piry, S., Alapetite, A., Cornuet, J. M., Paetkau, D., Baudouin, L., & Estoup, A. (2004). GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, *95*(6), 536–539. https://doi.org/10.1093/jhered/esh074

Pogson, G. H., & Fevolden, S. E. (2003). Natural selection and the genetic differentiation of

coastal and Arctic populations of the Atlantic cod in northern Norway: A test involving nucleotide sequence variation at the pantophysin (PanI) locus. *Molecular Ecology*, *12*(1), 63–74. https://doi.org/10.1046/j.1365-294X.2003.01713.x

Pomerantz, A., Peñafiel, N., Arteaga, A., Bustamante, L., Pichardo, F., Coloma, L. A., Barrio-Amorós, C. L., Salazar-Valenzuela, D., & Prost, S. (2018). Real-time DNA barcoding in a rainforest using nanopore sequencing: Opportunities for rapid biodiversity assessments and local capacity building. *GigaScience*, *7*(4), 1–14. https://doi.org/10.1093/gigascience/giy033

Poulsen, N. A., Hemmer-Hansen, J., Loeschcke, V., Carvalho, G. R., & Nielsen, E. E. (2011). Microgeographical population structure and adaptation in Atlantic cod Gadus morhua: Spatio-temporal insights from gene-associated DNA markers. *Marine Ecology Progress Series*, *436*, 231–243. https://doi.org/10.3354/meps09246

Prasetyo, A. P., McDevitt, A. D., Murray, J. M., Barry, J., Agung, F., Muttaqin, E., & Mariani, S. (2021). Shark and ray trade in and out of Indonesia: Addressing knowledge gaps on the path to sustainability. *Marine Policy*, *133*(December 2020), 104714. https://doi.org/10.1016/j.marpol.2021.104714

Qin, J., Vasefi, F., Hellberg, R. S., Akhbardeh, A., Isaacs, R. B., Yilmaz, A. G., Hwang, C., Baek, I., Schmidt, W. F., & Kim, M. S. (2020). Detection of fish fillet substitution and mislabeling using multimode hyperspectral imaging techniques. *Food Control*, *114*(March). https://doi.org/10.1016/j.foodcont.2020.107234

Quintrel, M., Jilberto, F., Sepúlveda, M., Marín, M. E., Véliz, D., Araneda, C., & Larraín, M. A. (2021). Development and validation of a multi-locus PCR-HRM method for species identification in mytilus genus with food authenticity purposes. *Foods*, *10*(8). https://doi.org/10.3390/foods10081684

R Core Development Team. (2018). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing.

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/

Ramos, R., & González-Solís, J. (2012). Trace me if you can: The use of intrinsic biogeochemical markers in marine top predators. *Frontiers in Ecology and the*

Environment, *10*(5), 258–266. https://doi.org/10.1890/110140

Rampazzo, F., Tosi, F., Tedeschi, P., Gion, C., Arcangeli, G., Brandolini, V., Giovanardi, O., Maietti, A., & Berto, D. (2020). Preliminary multi analytical approach to address geographic traceability at the intraspecific level in Scombridae family. *Isotopes in Environmental and Health Studies*, *56*(3), 260–279. https://doi.org/10.1080/10256016.2020.1739671

Ramstad, K. M., Woody, C. A., Sage, G. K., & Allendorf, F. W. (2004). Founding events influence genetic population structure of sockeye salmon (Oncorhynchus nerka) in Lake Clark, Alaska. *Molecular Ecology*, *13*(2), 277–290. https://doi.org/10.1046/j.1365-294X.2003.2062.x

Rannala B, & Mountain JL. (1997). Detecting immigration by using multilocus genotypes. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *94*(17), 9197–9201.

Rasmussen Hellberg, R. S., & Morrissey, M. T. (2011). Advances in DNA-based techniques for the detection of seafood species substitution on the commercial market. *Journal of Laboratory Automation*, *16*(4), 308–321. https://doi.org/10.1016/j.jala.2010.07.004

RedTractor. (2013). *BQAP Isotope Testing [Online]*. https://trade.redtractor.org.uk/licensing-information/traceability-challenges/bqap-isotope-testing

Reiss, H., Hoarau, G., Dickey-Collas, M., & Wolff, W. J. (2009). Genetic population structure of marine fish: Mismatch between biological and fisheries management units. *Fish and Fisheries*, *10*(4), 361–395. https://doi.org/10.1111/j.1467-2979.2008.00324.x

Rice, J. E., Reis, A. H., Rice, L. M., Carver-Brown, R. K., & Wangh, L. J. (2012). Fluorescent signatures for variable DNA sequences. *Nucleic Acids Research*, *40*(21), 1–10. https://doi.org/10.1093/nar/gks731

Riginos, C., Douglas, K. E., Jin, Y., Shanahan, D. F., & Treml, E. A. (2011). Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, *34*(4), 566–575. https://doi.org/10.1111/j.1600-0587.2010.06511.x

Rodriguez-Salvador, B., & Dopico, D. C. (2020). Understanding the value of traceability of

fishery products from a consumer perspective. *Food Control*, *112*(November 2019), 107142. https://doi.org/10.1016/j.foodcont.2020.107142

Rubinoff, D., & Holland, B. S. (2005). Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology*, *54*(6), 952–961. https://doi.org/10.1080/10635150500234674

Sarvas, T. H., & Fevolden, S. E. (2005). Pantophysin (Pan I) locus divergence between inshore v. offshore and northern v. southern populations of Atlantic cod in the north-east Atlantic. *Journal of Fish Biology*, *67*(2), 444–469. https://doi.org/10.1111/j.0022-1112.2005.00738.x

Schlegel, J., & Rupf, R. (2010). Attitudes towards potential animal flagship species in nature conservation: A survey among students of different educational institutions. *Journal for Nature Conservation*, *18*(4), 278–290. https://doi.org/10.1016/j.jnc.2009.12.002

Selkoe, K. A., Henzler, C. M., & Gaines, S. D. (2008). Seascape genetics and the spatial ecology of marine populations. *Fish and Fisheries*, *9*(4), 363–377. http://dx.doi.org/10.1111/j.1467-2979.2008.00300.x

Sellers, G. S., Di Muri, C., Gómez, A., & Hänfling, B. (2018). Mu-DNA: a modular universal DNA extraction method adaptable for a wide range of sample types. *Metabarcoding and Metagenomics*, *2*, e24556. https://doi.org/10.3897/mbmg.2.24556

Sembiring, A., Pertiwi, N. P. D., Mahardini, A., Wulandari, R., Kurniasih, E. M., Kuncoro, A. W., Cahyani, N. K. D., Anggoro, A. W., Ulfa, M., Madduppa, H., Carpenter, K. E., Barber, P. H., & Mahardika, G. N. (2015). DNA barcoding reveals targeted fisheries for endangered sharks in Indonesia. *Fisheries Research*, *164*, 130–134. https://doi.org/10.1016/j.fishres.2014.11.003

Shehata, H. R., Bourque, D., Steinke, D., Chen, S., & Hanner, R. (2019). Survey of mislabelling across finfish supply chain reveals mislabelling both outside and within Canada. *Food Research International*, *121*(February 2018), 723–729. https://doi.org/10.1016/j.foodres.2018.12.047

Shi, R., Xiong, X., Huang, M., Xu, W., Li, Y., Cao, M., & Xiong, X. (2020). High resolution melting (HRM) analysis of a 12S rRNA mini barcode as a novel approach for codfish

species authentication in processed fish products. *European Food Research and Technology*, *246*(5), 891–899. https://doi.org/10.1007/s00217-020-03456-5

Sick, K. (1965). Haemoglobin Polymorphism of Cod in the Baltic and the Danish Belt Sea. *Hereditas*, *54*(1), 19–48.

Silva, A. J., & Hellberg, R. S. (2021). DNA-based techniques for seafood species authentication. In *Advances in Food and Nutrition Research* (1st ed., Vol. 95). Elsevier Inc. https://doi.org/10.1016/bs.afnr.2020.09.001

Silva, G., Lima, F. P., Martel, P., & Castilho, R. (2014). Thermal adaptation and clinal mitochondrial DNA variation of European anchovy. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1792). https://doi.org/10.1098/rspb.2014.1093

Sirianni, N. M., Yuan, H., Rice, J. E., Kaufman, R. S., Deng, J., Fulton, C., & Wangh, L. J. (2016). Closed-Tube Barcoding. *Genome*, *59*(11), 1049–1061.

Skarstein, T. H., Westgaard, J. I., & Fevolden, S. E. (2007). Comparing microsatellite variation in north-east Atlantic cod (Gadus morhua L.) to genetic structuring as revealed by the pantophysin (Pan I) locus. *Journal of Fish Biology*, *70*(SUPPL. C), 271–290. https://doi.org/10.1111/j.1095-8649.2007.01456.x

Spencer, E. T., Richards, E., Steinwand, B., Clemons, J., Dahringer, J., Desai, P., Fisher, M., Fussell, S., Gorman, O., Jones, D., Le, A., Long, K., McMahan, C., Moscarito, C., Pelay, C., Price, E., Smith, A., VanSant, A., & Bruno, J. F. (2020). A high proportion of red snapper sold in North Carolina is mislabeled. *PeerJ*, *2020*(6), 1–10. https://doi.org/10.7717/peerj.9218

Springmann, M., Godfray, H. C. J., Rayner, M., & Scarborough, P. (2016). Analysis and valuation of the health and climate change cobenefits of dietary change. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(15), 4146–4151. https://doi.org/10.1073/pnas.1523119113

Stehfest, E., Bouwman, L., Van Vuuren, D. P., Den Elzen, M. G. J., Eickhout, B., & Kabat, P. (2009). Climate benefits of changing diet. *Climatic Change*, *95*(1–2), 83–102. https://doi.org/10.1007/s10584-008-9534-6

Stepien, C. A. (1999). Phylogeographical structure of the dover sole Microstomus pacificus: The larval retention hypothesis and genetic divergence along the deep

continental slope of the northeastern pacific ocean. *Molecular Ecology*, *8*(6), 923–939. https://doi.org/10.1046/j.1365-294X.1999.00643.x

Sumaila, U. R., Zeller, D., Hood, L., Palomares, M. L. D., Li, Y., & Pauly, D. (2020). Illicit trade in marine fish catch and its effects on ecosystems and people worldwide. *Science Advances*, *6*(9), 1–8. https://doi.org/10.1126/sciadv.aaz3801

Swartz, W., Rashid Sumaila, U., Watson, R., & Pauly, D. (2010). Sourcing seafood for the three major markets: The EU, Japan and the USA. *Marine Policy*, *34*(6), 1366–1373. https://doi.org/10.1016/j.marpol.2010.06.011

Tabrizi, A. (2021). *Seaspiracy*. Netflix.

Tagliavia, M., Nicosia, A., Salamone, M., Biondo, G., Bennici, C. D., Mazzola, S., & Cuttitta, A. (2016). Development of a fast DNA extraction method for sea food and marine species identification. *Food Chemistry*, *203*, 375–378. https://doi.org/10.1016/j.foodchem.2016.02.095

Thilsted, S. H., Thorne-Lyman, A., Webb, P., Bogard, J. R., Subasinghe, R., Phillips, M. J., & Allison, E. H. (2016). Sustaining healthy diets: The role of capture fisheries and aquaculture for improving nutrition in the post-2015 era. *Food Policy*, *61*, 126–131. https://doi.org/10.1016/j.foodpol.2016.02.005

Trueman, C. N., MacKenzie, K. M., & St John Glew, K. (2017). Stable isotope-based location in a shelf sea setting: accuracy and precision are comparable to light-based location methods. *Methods in Ecology and Evolution*, *8*, 232–240.

Trueman, Clive N., & St John Glew, K. (2019). Isotopic tracking of marine animal movement. In *Tracking Animal Migration with Stable Isotopes* (pp. 137–172). Elsevier Inc. https://doi.org/10.1016/B978-0-12-814723-8.00006-4

UN. (2019). *The Sustainable Development Goals Report 2019*. https://doi.org///doi.org/10.18356/55eb9109-en

Urbina, I. (2019). *The Outlaw Ocean: Journeys across the last untamed frontier*. Knopf.

Vandamme, S. G., Griffiths, A. M., Taylor, S.-A., Di Muri, C., Hankard, E. A., Towne, J. A., Watson, M., & Mariani, S. (2016). Sushi barcoding in the UK: another kettle of fish. *PeerJ*, *4*, e1891. https://doi.org/10.7717/peerj.1891

Vidgen, H. A., & Gallegos, D. (2014). Defining food literacy and its components. *Appetite*,

76, 50–59. https://doi.org/10.1016/j.appet.2014.01.010

Vileisis, A. (2008). *Kitchen literacy: how we lost knowledge of where food comes from and why we need to get it back*. Washington, DC: Island Press.

Viñas, J., Sanz, N., Peñarrubia, L., García-Marín, J.-L., Roldán, M.-I., & Pla, C. (2014). Genetic population structure of European anchovy in the Mediterranean Sea and the Northeast Atlantic Ocean using sequence analysis of the mitochondrial DNA control region. *Journal of Marine Science*, *71*(2), 391–397. https://doi.org/10.1038/278097a0

Vitale, S., Biondo, F., Giosuè, C., Bono, G., Okpala, C. O. R., Piazza, I., Sprovieri, M., & Pipitone, V. (2020). Consumers' perception and willingness to pay for eco-labeled seafood in Italian hypermarkets. *Sustainability (Switzerland)*, *12*(4), 1–13. https://doi.org/10.3390/su12041434

Waples, R. S. (1987). A multispecies approach to the analysis of gene flow in marine shore fishes. *Evolution*, *41*(2), 385–400. https://doi.org/10.1111/j.1558-5646.1987.tb05805.x

Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, *15*(6), 1419–1439. https://doi.org/10.1111/j.1365-294X.2006.02890.x

Ward, R. D., Hanner, R., & Hebert, P. D. N. (2009). The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology*, *74*(2), 329–356. https://doi.org/10.1111/j.1095-8649.2008.02080.x

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1462), 1847–1857. https://doi.org/10.1098/rstb.2005.1716

Warner, K., Timme, W., & Lowell, B. (2012). Widespread Seafood Fraud Found in New York City. In *Oceana* (Issue December).

Warnes, R. G., Bolker, B., Lumley, T., & Johnson, R. C. (2018). *gmodels: Various R Programming Tools for Model Fitting. R package version 2.18.1*. https://cran.r-project.org/package=gmodels

Weersing, K., & Toonen, R. J. (2009). Population genetics, larval dispersal, and

connectivity in marine systems. *Marine Ecology Progress Series*, *393*, 1–12. https://doi.org/10.3354/meps08287

Wennevik, V., Quintela, M., Skaala, Ø., Verspoor, E., Prusov, S., & Glover, K. A. (2019). Population genetic analysis reveals a geographically limited transition zone between two genetically distinct Atlantic salmon lineages in Norway. *Ecology and Evolution*, *9*(12), 6901–6921. https://doi.org/10.1002/ece3.5258

Wickham, H. (2016). ggplot2: elegant graphics for data analysis. *Springer*.

Willette, D. A., Allendorf, F. W., Barber, P. H., Barshis, D. J., Carpenter, K. E., Crandall, E. D., Cresko, W. A., Fernandez-Silva, I., Matz, M. V., Meyer, E., Santos, M. D., Seeb, L. W., & Seeb, J. E. (2014). So, you want to use next-generation sequencing in marine systems? Insight from the Pan-Pacific Advanced Studies Institute. *Bulletin of Marine Science*, *90*(1), 79–122. https://doi.org/10.5343/bms.2013.1008

Williams, M., Hernandez-Jover, M., & Shamsi, S. (2020). Fish substitutions which may increase human health risks from zoonotic seafood borne parasites: A review. *Food Control*, *118*(June), 107429. https://doi.org/10.1016/j.foodcont.2020.107429

Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., Jackson, J. B., Lotze, H. K., Micheli, F., Palumbi, S. R., & Others, A. (2006). Impacts of biodiversity loss on ocean ecosystem services. *Science*, *314*(5800), 787–791. https://doi.org/10.1126/science.1137946

Worm, B., Hilborn, R., Baum, J. K., Branch, T. A., Collie, J. S., Costello, C., Fogarty, M. J., Fulton, E. A., Hutchings, J. A., Jennings, S., Jensen, O. P., Lotze, H. K., Mace, P. M., McClanahan, T. R., Minto, C., Palumbi, S. R., Parma, A. M., Ricard, D., Rosenberg, A. A., … Zeller, D. (2009). Rebuilding Global Fisheries. *Science*, *325*(5940), 578–585. https://doi.org/10.1126/science.1173146

Wright, S. (1931). Evolution in mendelian populations. *Genetics*, *16*(97). https://doi.org/10.1007/BF02459575

Wright, S. (1965). The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating. *Evolution*, *19*(3), 395. https://doi.org/10.2307/2406450

Xiong, X., Guardone, L., Cornax, M. J., Tinacci, L., Guidi, A., Gianfaldoni, D., & Armani, A. (2016). DNA barcoding reveals substitution of Sablefish (Anoplopoma fimbria) with

Patagonian and Antarctic Toothfish (Dissostichus eleginoides and Dissostichus mawsoni) in online market in China: How mislabeling opens door to IUU fishing. *Food Control*, *70*, 380–391. https://doi.org/10.1016/j.foodcont.2016.06.010

Yancy, H. F., Zemlak, T. S., Mason, J. A., Washington, J. D., Tenge, B. J., Nguyen, N. L. T., Barnett, J. D., Savary, W. E., Hill, W. E., Moore, M. M., Fry, F. S., Randolph, S. C., Rogers, P. L., & Hebert, P. D. N. (2008). Potential use of DNA barcodes in regulatory science: Applications of the regulatory fish encyclopedia. *Journal of Food Protection*, *71*(1), 210–217. https://doi.org/10.4315/0362-028X-71.1.210

Yasuda, T., & Bowen, R. E. (2006). Chain of custody as an organizing framework in seafood risk reduction. *Marine Pollution Bulletin*, *53*, 640–649.

Young, E. F., Belchier, M., Hauser, L., Horsburgh, G. J., Meredith, M. P., Murphy, E. J., Pascoal, S., Rock, J., Tysklind, N., & Carvalho, G. R. (2015). Oceanography and life history predict contrasting genetic population structure in two Antarctic fish species. *Evolutionary Applications*, *8*(5), 486–509. https://doi.org/10.1111/eva.12259

Zarraonaindia, I., Iriondo, M., Albaina, A., Pardo, M. A., Manzano, C., Grant, W. S., Irigoien, X., & Estonba, A. (2012). Multiple SNP markers reveal fine-scale population and deep phylogeographic structure in european anchovy (Engraulis encrasicolus L.). *PLoS ONE*, *7*(7). https://doi.org/10.1371/journal.pone.0042201

Zeller, D., & Pauly, D. (2019). Viewpoint: Back to the future for fisheries, where will we choose to go? *Global Sustainability*, *2*. https://doi.org/10.1017/sus.2019.8

Zhang, X., Cheng, J., Han, D., Zhao, X., Chen, X., & Liu, Y. (2019). Geographical origin traceability and species identification of three scallops (Patinopecten yessoensis, Chlamys farreri, and Argopecten irradians) using stable isotope analysis. *Food Chemistry*, *299*.

Supplementary Material

Supplementary Material - Chapter 4

Table S4.1 Code name, scientific name, and common name of the finfish species for which life-history trait analysis was conducted.

CODE	Species scientific name	Species common name
ALB	*Thunnus alalunga*	Albacore Tuna
ALK	*Gadus chalcogrammus*	Pollock (walleye)
ALP	*Pleuronectes quadrituberculatus*	Alaskan Plaice
ANA	*Engraulis anchoita*	Anchovy (Argentine)
ANE	*Engraulis encrasicolus*	Anchovy (European)
ANI	*Champsocephalus gunnari*	Mackerel icefish
ARF	*Atheresthes stomias*	Flounder (arrow-tooth)
ARU	*Argentina silus*	Smelt (silver)
BFT	*Thunnus thynnus*	Atlantic Bluefin Tuna
BLI	*Molva dypterygia*	Blue Ling
CAA	*Anarhichas lupus*	Wolffish (Atlantic)
CHP	*Sardinops sagax*	Sardine
CLI	*Ophiodon elongatus*	Lingcod
COD	*Gadus morhua*	Cod (Atlantic)
DGS	*Squalus acanthias*	Spiny Dogfish
FTS	*Hippoglossoides elassodon*	Sole (flathead)
GHL	*Reinhardtius hippoglossoides*	Greenland Halibut
GLZ	*Glyptocephalus zachirus*	Sole (rex)
HAD	*Melanogrammus aeglefinus*	Haddock
HAL	*Hippoglossus hippoglossus*	Halibut (Atlantic)
HAP	*Hippoglossus stenolepis*	Halibut (Pacific)
HER	*Clupea harengus*	Herring
HKC	*Merluccius paradoxus*	Cape Hake

HKE	*Merluccius merluccius*	Hake (European)
JAH	*Raja rhina*	Longnose Skate
KAF	*Atheresthes evermanni*	Flounder (Kamchatka)
KOB	*Argyrosomus hololepidotus*	Mulloway
LEM	*Microstomus kitt*	Lemon Sole
LIN	*Molva molva*	Ling
LUM	*Cyclopterus lumpus*	Lumpfish(=Lumpsucker)
MAC	*Scomber scombrus*	Mackerel
MAD	*Aldrichetta forsteri*	Mullet (yellow-eye)
MEG	*Lepidorhombus whiffiagonis*	Megrim
MHA	*Brevoortia tyrannus*	Atlantic menhaden
MHG	*Brevoortia patronus*	Gulf menhaden
MIP	*Microstomus pacificus*	Dover Sole
MUF	*Mugil cephalus*	Mullet (flathead grey)
NHA	*Merluccius productus*	Hake (North Pacific)
NOP	*Trisopterus esmarkii*	Norway Pout
OPP	*Sebastes alutus*	Pacific Ocean Perch
ORY	*Hoplostethus atlanticus*	Orange roughy
PCO	*Gadus macrocephalus*	Cod (Pacific)
PIL	*Sardina pilchardus*	European Pilchard (Sardine)
PLE	*Pleuronectes platessa*	European Plaice
PLM	*Plectropomus maculatus*	Spotted Coral Grouper
POK	*Pollachius virens*	Saithe
POS	*Micromesistius australis*	Whiting (southern blue)
QLH	*Ammodytes marinus*	Lesser sand-eel
REN	*Sebastes fasciatus*	Acadian redfish
RFC	*Sebastes crameri*	Darkblotched rockfish
RFE	*Pleuronectes vetulus*	English Sole
RYO	*Sebastes flavidus*	Yellowtail rockfish
SAB	*Anoplopoma fimbria*	Sablefish

SBC	*Sebastes paucispinis*	Bocaccio rockfish
SFD	*Sebastes diploproa*	Splitnose rosckfish
SGO	*Sebastes goodei*	Chillipepper rockfish
SJU	*Sebastolobus alascanus*	Shortspine Thornyhead
SJZ	*Sebastolobus altivelis*	Longspine Thornyhead
SKJ	*Katsuwonus pelamis*	Skipjack Tuna
SOL	*Solea solea*	Sole
SPR	*Sprattus sprattus*	European Sprat
SWO	*Xiphias gladius*	Swordfish
THP	*Opisthonema libertate*	Pacific thread herring
TOA	*Dissostichus mawsoni*	Toothfish (Antarctic)
TOP	*Dissostichus eleginoides*	Toothfish (Patagonian)
USK	*Brosme brosme*	Tusk (=Cusk)
WHB	*Micromesistius poutassou*	Blue Whiting
WRO	*Sebastes entomelas*	Widow rockfish
YFT	*Thunnus albacares*	Yellowfin Tuna
ZAZ	*Lepidopsetta polyxystra*	Sole (Northern rock)

Supplementary Material - Chapter 5

Table S5.1 Metadata for the Atlantic cod samples.

Sample	Country sampled	City sampled	Year sampled	SM/FM	Market name	MSC (Y/N)	Claimed geographical catch area
1	FR	Paris	2019	FM	Marché Edgard Quinet	N	France
2	FR	Paris	2019	FM	Marché Edgard Quinet	N	Nomandie/Bretagne
3	FR	Paris	2019	FM	Marché Edgard Quinet	N	Nomandie/Bretagne
4	FR	Paris	2019	FM	Poissonnerie du Dome	N	Mer du Nord
6	FR	Paris	2019	SM	Monoprix	Y	Atlantic NE - Mer de Barents, Mer de Norvège, Islande, et Féroé
7	FR	Paris	2019	FM	Monoprix	N	Area 27 - V (Islande)
8	FR	Paris	2019	SM	Casino	Y	Mer Baltique, Islande, & Féroé
10	FR	St. Malo	2019	FM	Poissonnerie Gilles Guinemer	N	Manche & Mer Celtique
11	FR	St. Malo	2019	FM	Compagnie des pêches	N	Mer de Barents, Norvège, Spitzberg, ile aux ours
12	FR	Guilvinec	2019	FM	Marée du jour	N	Zone 27 - VII - Manche et Mer Celtique
14	FR	Concarneau	2019	FM	Poissonnerie La Roche	N	Manche & Mer Celtique
15	FR	Quimper	2019	FM	Poissonnier Marc	N	written: Mer du Nord (said: Manche & Mer Celtique)
16	FR	Quimper	2019	FM	Poissonnerie Salaun	N	Golf de Gascogne
17	FR	Paris	2019	SM	La Vie Claire	Y	Atlantic NE - FAO27 - Islande
18	FR	Paris	2019	SM	Bio c'Bon	Y	Atlantic NE - FAO27 - Islande
19	FR	Paris	2019	SM	Bio c'Bon	Y	Atlantic NE - FAO27 - Islande
20	FR	Paris	2019	SM	Monoprix Sebastopol	Y	Islande et Féroé
21	FR	Paris	2019	SM	Monoprix Sebastopol	Y	Mer de Norvège, Islande, et Féroé
22	FR	Paris	2019	SM	LiDL	Y	Atlantic NE - Islande et Féroé
23	FR	Paris	2019	SM	G20	N	Atlantic NE - Mer de Barents et Mer de Norvège
24	FR	Paris	2019	SM	Monoprix	N	Atlantic Nord OUEST
25	FR	Paris	2019	SM	Monoprix	N	Atlantic Nord OUEST
26	FR	Paris	2019	SM	Monoprix	N	Atlantic NE - Mer de Barents, Norvège, Ile aux ours, Spitzerg, Feroé, Mer du Nord
27	FR	Paris	2019	SM	Uexpress	N	Atlantic NE - Mer de Norvège, Islande, et Féroé
28	FR	Paris	2019	FM	Marché Raspail	Y	FAO 27 - Va - Iceland
29	UK	Newcastle	2019	FM	JR Fisheries	N	North Sea
30	UK	Liverpool	2020	SM	TESCO	Y	Norwegian Sea
31	UK	Manchester	2020	SM	Morrisons	N	North-West & North-East Atlantic; Sub areas I, II, & V
32	UK	Manchester	2020	SM	Morrisons	N	North-West & North-East Atlantic; Sub areas I, II, & V

33	UK	Manchester	2020	SM	COOP	Y	Norwegian Sea, Barents Sea, Iceland & Faroes Grounds (NEA 27 subareas I, II, V)
34	UK	Manchester	2020	SM	Morrisons	N	Norwegian Sea, Icelandic Grounds, Barents Sea & North-West Atlantic
35	UK	Manchester	2020	SM	Morrisons	N	North-West & North-East Atlantic; Sub areas I, II, & V
36	UK	Manchester	2020	FM	Out of the Blue	Y	Rockall/Shetland
37	UK	Manchester	2020	SM	COOP	Y	NEA: Norwegian sea, spitzbergen, bear island, barents sea, iceland, faroe
38	UK	Manchester	2020	SM	Morissons	N	NEA: Norwegian sea, Iceland grounds, Barents sea, North West Atlantic
39	UK	Manchester	2020	SM	SPAR	N	Not listed
40	UK	Manchester	2020	SM	Aldi Arndale	Y	FAO ia, ib, iia, iib, Barents sea, Norwegian sea, Spitzbergen, Bear Isd
41	UK	Manchester	2020	SM	Aldi Arndale	Y	North EA (27), Sub: I-Barents sea, II-Norwegian sea, Spitzbergen, Bear Island, V-Iceland & Feroe
42	UK	Manchester	2020	SM	Tesco Extra	Y	NEA Atlantic: Norwegian sea
43	UK	Manchester	2020	SM	Tesco Extra	Y	NEA Atlantic: Norwegian sea
44	UK	Manchester	2020	SM	Tesco Extra	Y	NEA Atlantic: Barents sea
45	UK	Manchester	2020	SM	Tesco Extra	Y	NEA Atlantic: Barents sea
46	UK	Manchester	2020	SM	LIDL	Y	NA: Subareas Ia-b, 2a-b, 5a
47	UK	Manchester	2020	SM	LIDL	Y	NA: Subareas Ia-b, 2a-b, 5a
48	UK	Manchester	2020	SM	M&S	N	North-East Atlantic & North West Atlantic
49	UK	Manchester	2020	SM	M&S	N	NEA: Icelandic Grounds
50	UK	Manchester	2020	SM	M&S	N	NEA: Icelandic Grounds
51	UK	Manchester	2020	SM	Iceland Stores	Y	NEA: Norwegian sea, Spitzbergen, Bear Island, Barents sea
52	UK	Manchester	2020	SM	Iceland Stores	Y	NEA: Norwegian sea, Spitzbergen, Bear Island, Barents sea
53	UK	Manchester	2020	SM	Iceland Stores	Y	FAO 27: Barents sea, Norwegian sea, Spitzbergen, Bear Island
54	UK	Manchester	2020	SM	Aldi	Y	NEA: I, ii, v - Barents sea, Norwegian sea, Spitzbergen, Bear island, Iceland & Faroe
55	UK	Manchester	2020	SM	Aldi	Y	NEA: I, ii - Barents sea, Norwegian sea, Spitzbergen, Bear island
56	UK	Manchester	2020	SM	Aldi	Y	NEA: I, ii - Barents sea, Norwegian sea, Spitzbergen, Bear island
57	UK	Manchester	2020	SM	Aldi	Y	NEA: I, ii - Barents sea, Norwegian sea, Spitzbergen, Bear island
58	UK	Manchester	2020	FM	Direct Fisheries	N	Cornwall
59	UK	Manchester	2020	FM	Direct Fisheries	N	Cornwall
60	UK	Manchester	2020	FM	Direct Fisheries	N	Cornwall
61	UK	Manchester	2020	FM	Direct Fisheries	N	Cornwall
62	UK	Manchester	2020	FM	Direct Fisheries	N	Cornwall
63	UK	Manchester	2020	FM	Direct Fisheries	N	Cornwall
64	UK	Manchester	2020	FM	Sea it Fresh	N	Scottish - Landed in Peterhead
65	UK	Manchester	2020	FM	Sale Fish&Seafoods	N	West Scotland - North Iceland
66	UK	NA	2020	FM	Roaming fishmonger	N	Aberdeen

67	UK	Fraserburgh	2020	FM	The Fish Plaice	N	North Sea - Area 27
68	UK	Fraserburgh	2020	FM	The Fish Plaice	N	North Sea - Area 27
69	UK	Fraserburgh	2020	FM	Jack Taylor	N	North Sea
70	UK	Stone haven	2020	FM	The creel fishmonger	N	Landed in Peterhead - Caught North Sea
71	UK	Gourdon	2020	FM	Gourdine	N	Landed in Gourdon - Caught offshore (North Sea)
72	UK	Arbroath	2020	FM	C Lyons Fresh Fish	N	Landed in Peterhead
73	UK	Arbroath	2020	FM	E&O Fish Arbroath	N	Landed in Peterhead (North Sea)
74	UK	Arbroath	2020	FM	The Fish Hoose	N	Landed in Peterhead (North Sea)
75	UK	Edinburgh	2020	FM	G. Armstrong	N	Scottish cod (North Sea)
76	UK	Edinburgh	2020	FM	J. Williamson Fishmonger	N	Landed in Peterhead (North Sea)
77	UK	Edinburgh	2020	FM	Welch Fishmonger	N	North Sea
78	UK	Manchester	2013	SM	NA	NA	North-East Atlantic
79	UK	Manchester	2013	FM	NA	NA	North-East Atlantic
80	UK	Manchester	2013	FM	NA	NA	North-East Atlantic
81	UK	Manchester	2013	FM	NA	NA	North-East Atlantic
82	UK	Manchester	2013	FM	NA	NA	North-East Atlantic
83	UK	Manchester	2013	SM	NA	NA	North-East Atlantic
84	UK	Manchester	2013	SM	NA	NA	North-East Atlantic
85	UK	Manchester	2013	SM	NA	NA	North-East & North-West
86	UK	Manchester	2013	SM	NA	NA	North-East Atlantic
87	UK	Manchester	2013	FM	NA	NA	North-East Atlantic
88	UK	Manchester	2013	FM	NA	NA	North-East Atlantic
89	UK	Manchester	2013	SM	NA	NA	North-East Atlantic
90	UK	Manchester	2013	SM	NA	NA	North-East Atlantic
91	UK	Glasgow	2013	SM	NA	NA	Atlantic
92	UK	Glasgow	2013	SM	NA	NA	Iceland
93	UK	Glasgow	2013	SM	NA	NA	North-East & North-West
94	UK	Glasgow	2013	SM	NA	NA	North-East Atlantic
95	UK	Glasgow	2013	SM	NA	NA	North-East Atlantic
96	UK	Glasgow	2013	FM	NA	NA	North-East Atlantic
97	UK	Glasgow	2013	FM	NA	NA	Atlantic
98	UK	Glasgow	2013	SM	NA	NA	Atlantic
99	UK	Glasgow	2013	SM	NA	NA	Atlantic
100	GE	Kiel	2020	SM	Citti Markt Kiel	N	FAO27, Norwegian Sea
101	GE	Kiel	2020	SM	Citti Markt Kiel	Y	on package no information; Trackingcode: FAO27, Norwegian Sea
102	GE	Kiel	2020	SM	Citti Markt Kiel	Y	FAO27; Barents Sea, Norwegian Sea, Spitsbergen and Bear Island
103	GE	Kiel	2020	SM	Citti Markt Kiel	Y	FAO27; Barents Sea/Spitsbergen and Bear Island
104	GE	Kiel	2020	SM	Citti Markt Kiel	Y	FAO27; Island
105	GE	Kiel	2020	SM	Citti Markt Kiel	N	FAO27; Norwegian Sea
106	GE	Kiel	2020	SM	Citti Markt Kiel	N	FAO27; North Sea
107	GE	Kiel	2020	FM	Matjes Lange	N	FAO27; North Sea
108	GE	Kiel	2020	FM	Keste	N	FAO27; Lofoten, Norway
109	GE	Kiel	2020	SM	Famila Kiel	N	FAO27; Northern North Sea
110	GE	Hamburg	2020	SM	Mercado, HH-Altona	Y	FAO27; Barents Sea, Norwegian Sea, Spitsbergen and Bear Island
111	GE	Hamburg	2020	FM	HH-Altona	N	North Antlantic; Lofoten ,Norway
112	GE	Hamburg	2020	FM	Mercado, HH-Altona	N	Nordsee FAO (?)

113	GE	Hamburg	2020	SM	Bahrenfelder Straße	Y	FAO 27; Barents Sea or Norwegian Sea
114	GE	Hamburg	2020	SM	Brauerknachtgraben	Y	North East Atlantic, Norwegian Sea
115	GE	Hamburg	2020	SM	Brauerknachtgraben	Y	North East Atlantic, Island
116	GE	Hamburg	2020	SM	Metro	Y	FAO27; Barents Sea
117	GE	Hamburg	2020	SM	Metro	Y	FAO27, Norwegian Sea
118	GE	Hamburg	2020	FM	Mein Fischladen	N	FAO27, Norwegian Sea
119	GE	Hamburg	2020	SM	Frischeparadies	N	FAO27; Iceland
120	SP	Vigo	2020	SM	ALCAMPO	N	FAO 27 II / FAO 27 V
121	SP	Vigo	2020	SM	ALCAMPO	N	FAO 27 II b
122	SP	Vigo	2020	SM	ALCAMPO	N	fao 27 Norway
123	SP	Vigo	2020	SM	ALCAMPO	Y	FAO 27 Mar de Noruega, Spitzberg e Islas de los Osos
124	SP	Vigo	2020	SM	CARREFOUR	Y	FAO 27/IIa2. Northeast Atlantic and Norweian Sea
125	SP	Vigo	2020	SM	CARREFOUR	N	Atlántico NE, Subzona Va (Islandia)
126	SP	Vigo	2020	SM	CARREFOUR	Y	Mar de Noruega, Spitzberg e Isla de osos. Islandia e Islas Feroe.
127	SP	Vigo	2020	SM	CARREFOUR	N	Atlántico NE. Islandia. Puede venir de otras zonas.
128	SP	Vigo	2020	SM	CARREFOUR	Y	FAO 27/II. Mar de Noruega
129	SP	Vigo	2020	FM	JULIA QUINTEIRO	Y	FAO 27/IIa2. Norweian Sea.
130	SP	Vigo	2020	FM	PESCADOS TROULO	N	FAO 27. Atlantic northeast (Galician waters)
131	SP	Vigo	2020	FM	PESCADOS TROULO	N	FAO 27. Atlantic northeast (Galician waters)
132	SP	Vigo	2020	FM	PESCADOS TROULO	N	FAO 27. Atlantic northeast (Galician waters)
133	Barents sea	NA	2020	Voucher Samples University of Oslo			
134	Barents sea	NA	2020	Voucher Samples University of Oslo			
135	Barents sea	NA	2020	Voucher Samples University of Oslo			
136	Barents sea	NA	2020	Voucher Samples University of Oslo			
137	Barents sea	NA	2020	Voucher Samples University of Oslo			
138	Barents sea	NA	2020	Voucher Samples University of Oslo			
139	Barents sea	NA	2020	Voucher Samples University of Oslo			
140	Barents sea	NA	2020	Voucher Samples University of Oslo			
141	Barents sea	NA	2020	Voucher Samples University of Oslo			
142	Barents sea	NA	2020	Voucher Samples University of Oslo			
143	Svalbard	NA	2020	Voucher Samples Norwegian Polar Institute			
144	Svalbard	NA	2020	Voucher Samples Norwegian Polar Institute			
145	Svalbard	NA	2020	Voucher Samples Norwegian Polar Institute			
146	Svalbard	NA	2020	Voucher Samples Norwegian Polar Institute			
147	Svalbard	NA	2020	Voucher Samples Norwegian Polar Institute			
148	Svalbard	NA	2020	Voucher Samples Norwegian Polar Institute			
149	Svalbard	NA	2020	Voucher Samples Norwegian Polar Institute			

150	Svalbard	NA	2020	Voucher Samples Norwegian Polar Institute
151	Svalbard	NA	2020	Voucher Samples Norwegian Polar Institute
152	Svalbard	NA	2020	Voucher Samples Norwegian Polar Institute
153	North sea	NA	2020	Voucher Samples Cefas, UK
154	North sea	NA	2020	Voucher Samples Cefas, UK
155	North sea	NA	2020	Voucher Samples Cefas, UK
156	North sea	NA	2020	Voucher Samples Cefas, UK
157	North sea	NA	2020	Voucher Samples Cefas, UK
158	North sea	NA	2020	Voucher Samples Cefas, UK
159	North sea	NA	2020	Voucher Samples Cefas, UK
160	North sea	NA	2020	Voucher Samples Cefas, UK
161	North sea	NA	2020	Voucher Samples Cefas, UK
162	North sea	NA	2020	Voucher Samples Cefas, UK
163	UK	NA	2019	

Sample	FASTFISH-ID	Amplified loci	Primary Assignment	Assignment score (%)	Paetkau NS probability	Paetkau NEA probability
1	*G. morhua*	9	NEA	97.813	0.038	0.23
2	*G. morhua*	9	NS	83.77	0.437	0.284
3	*G. morhua*	9	NEA	98.048	0.147	0.646
4	*G. morhua*	9	NEA	96.029	0.05	0.238
6	*G. morhua*	9	NEA	96.711	0.117	0.429
7	*G. morhua*	7	NS	80.561	0.157	0.046
8	*G. morhua*	9	NEA	78.015	0.218	0.421
10	*G. morhua*	8	NEA	56.59	0.136	0.198
11	*G. morhua*	9	NEA	92.349	0.054	0.205
12	*G. morhua*	7	NEA	97.62	0.055	0.252
14	*G. morhua*	9	NEA	75.534	0.025	0.031
15	*G. morhua*	8	NEA	96.785	0.049	0.211
16	*G. morhua*	9	NEA	88.254	0.053	0.172
17	*G. morhua*	9	NEA	95.424	0.056	0.248
18	*G. morhua*	9	NEA	93.462	0.038	0.16
19	*G. morhua*	8	NEA	66.006	0.336	0.459
20	*G. morhua*	9	NS	69.781	0.362	0.308
21	*G. morhua*	9	NS	52.546	0.344	0.387
22	*G. morhua*	9	NS	98.729	0.859	0.222
23	*G. morhua*	7	NEA	93.926	0.117	0.369
24	*G. morhua*	9	NEA	91.311	0.115	0.294
25	*G. morhua*	9	NEA	89.801	0.025	0.061
26	*G. morhua*	9	NEA	94.355	0.106	0.317
27	*G. morhua*	9	NEA	99.775	0.045	0.574

28	*G. morhua*	8	NEA	87.476	0.094	0.229
29	*G. morhua*	9	NS	98.259	0.129	0.012
30	*G. morhua*	9	NS	86.476	0.13	0.055
31	*G. morhua*	8	NS	59.66	0.064	0.061
32	*G. morhua*	9	NEA	62.77	0.159	0.242
33	*G. morhua*	8	NEA	99.629	0.038	0.504
34	*G. morhua*	9	NEA	94.56	0.143	0.451
35	*G. morhua*	9	NEA	54.783	0.109	0.133
36	*G. morhua*	9	NS	99.787	0.203	0.007
37	*G. morhua*	8	NEA	96.663	0.087	0.406
38	*G. morhua*	8	NEA	89.32	0.03	0.132
39	*G. morhua*	9	NEA	54.372	0.266	0.328
40	*G. morhua*	9	NEA	82.267	0.125	0.254
41	*G. morhua*	8	NEA	86.419	0.162	0.373
42	*G. morhua*	9	NEA	93.829	0.056	0.222
43	*G. morhua*	9	NEA	88.292	0.155	0.359
44	*G. morhua*	9	NEA	97.453	0.019	0.07
45	*G. morhua*	9	NS	77.142	0.289	0.235
46	*G. morhua*	9	NEA	97.754	0.06	0.31
47	*G. morhua*	9	NEA	88.168	0.056	0.181
48	*G. morhua*	9	NS	84.349	0.106	0.041
49	*G. morhua*	9	NEA	86.371	0.197	0.462
50	*G. morhua*	9	NEA	84.16	0.055	0.145
51	*G. morhua*	9	NEA	76.227	0.194	0.345
52	*G. morhua*	9	NEA	91.617	0.154	0.404
53	*G. morhua*	8	NS	66.758	0.519	0.386
54	*G. morhua*	9	NEA	96.491	0.038	0.212
55	*G. morhua*	8	NS	74.334	0.16	0.143
56	*G. morhua*	9	NS	82.438	0.392	0.266
57	*G. morhua*	9	NEA	62.77	0.159	0.242
58	*G. morhua*	9	NEA	86.736	0.025	0.053
59	*G. morhua*	8	NS	62.447	0.032	0.019
60	*G. morhua*	9	NEA	82.565	0.087	0.204
61	*G. morhua*	8	NS	99.71	0.32	0.015
62	*G. morhua*	9	NS	99.754	0.176	0.007
63	*G. morhua*	9	NS	53.729	0.064	0.059
64	*G. morhua*	8	NEA	99.501	0.072	0.772
65	*G. morhua*	9	NS	98.165	0.521	0.146
66	*M. aeglefinnus*	1	NEA	61.57	0.063	0.083
67	*G. morhua*	9	NEA	74.524	0.038	0.057
68	*G. morhua*	8	NEA	74.931	0.269	0.47
69	*G. morhua*	9	NEA	66.078	0.127	0.205
70	*G. morhua*	9	NS	75.762	0.123	0.07
71	*G. morhua*	8	NEA	54.12	0.166	0.246
72	*G. morhua*	9	NEA	54.494	0.199	0.273
73	*G. morhua*	9	NS	54.077	0.134	0.155

74	*M. aeglefinnus*	1	NEA	61.57	0.066	0.041
75	*G. morhua*	9	NS	95.076	0.223	0.079
76	*G. morhua*	9	NEA	98.299	0.034	0.234
77	*G. morhua*	9	NEA	53.943	0.045	0.049
78	*G. morhua*	9	NEA	76.032	0.136	0.247
79	*G. morhua*	9	NS	66.688	0.197	0.213
80	Unknown	9	NEA	79.248	0.151	0.283
81	*G. morhua*	9	NS	53.739	0.085	0.081
82	*G. morhua*	9	NEA	96.711	0.117	0.429
83	*G. morhua*	9	NS	58.854	0.198	0.232
84	*G. morhua*	9	NS	97.907	0.159	0.025
85	Unknown	8	NEA	73.749	0.17	0.318
86	*G. morhua*	5	NS	95.205	0.511	0.05
87	*G. morhua*	9	NEA	66.951	0.163	0.258
88	*G. morhua*	9	NEA	92.756	0.053	0.206
89	*G. morhua*	9	NEA	54.106	0.17	0.234
90	*G. morhua*	9	NEA	96.711	0.117	0.429
91	Unknown	4	NEA	99.07	0.039	0.438
92	Unknown	8	NEA	94.845	0.022	0.154
93	*G. morhua*	9	NEA	94.355	0.106	0.317
94	*G. morhua*	9	NS	71.036	0.219	0.22
95	*G. morhua*	8	NS	74.084	0.719	0.403
96	*G. morhua*	9	NS	85.699	0.051	0.017
97	Unknown	7	NEA	98.592	0.033	0.232
98	*G. macrocephalus*	6	NEA	99.987	0	0.095
99	*M. aeglefinnus*	3	NS	74.551	0.16	0.025
100	*G. morhua*	8	NS	58.085	0.484	0.456
101	*G. morhua*	9	NEA	95.339	0.064	0.258
102	*G. morhua*	8	NEA	96.663	0.088	0.368
103	*G. morhua*	9	NEA	96.792	0.029	0.181
104	*G. morhua*	9	NEA	86.595	0.179	0.425
105	*G. morhua*	9	NEA	73.603	0.045	0.086
106	*G. morhua*	8	NEA	88.418	0.064	0.197
107	*G. morhua*	9	NS	85.405	0.178	0.118
108	*G. morhua*	8	NEA	83.126	0.06	0.192
109	*G. morhua*	9	NEA	65.141	0.093	0.137
110	*G. morhua*	9	NS	99.493	0.53	0.069
111	*G. morhua*	9	NEA	91.617	0.154	0.404
112	*G. morhua*	9	NEA	94.46	0.156	0.494
113	*G. morhua*	9	NEA	91.812	0.042	0.171
114	Low amplification	9	NEA	69.707	0.277	0.438
115	*G. morhua*	9	NEA	50.975	0.305	0.357
116	*G. morhua*	9	NEA	82.134	0.134	0.264
117	*G. morhua*	9	NEA	97.556	0.165	0.711
118	*G. morhua*	9	NEA	73.669	0.215	0.377
119	*G. morhua*	9	NEA	95.197	0.103	0.33
120	*G. morhua*	8	NEA	97.864	0.038	0.27
121	*G. morhua*	9	NEA	81.85	0.143	0.284
122	*G. morhua*	9	NEA	88.168	0.056	0.181
123	*G. morhua*	9	NEA	57.783	0.159	0.225
124	*G. morhua*	9	NEA	85.798	0.143	0.304
125	*G. morhua*	9	NEA	96.316	0.016	0.049
126	Low amplification	9	NEA	89.298	0.097	0.252
127	*G. morhua*	9	NEA	97.141	0.059	0.285

128	*G. morhua*	9	NEA	88.092	0.163	0.393
129	*G. morhua*	9	NEA	94.052	0.092	0.284
130	*G. morhua*	9	NEA	54.017	0.218	0.287
131	*G. morhua*	8	NS	90.045	0.156	0.074
132	*G. morhua*	9	NS	92.888	0.342	0.188
133	Low amplification	9	NEA	97.614	0.038	0.231
134	*G. morhua*	9	NEA	96.65	0.127	0.469
135	Low amplification	8	NEA	62.728	0.138	0.227
136	Low amplification	9	NEA	91.311	0.115	0.294
137	Unknown	9	NEA	96.65	0.127	0.469
138	Unknown	9	NEA	94.46	0.156	0.494
139	Unknown	9	NEA	96.792	0.029	0.181
140	Low amplification	9	NEA	86.595	0.179	0.425
141	Low amplification	8	NEA	94.482	0.127	0.415
142	*G. morhua*	9	NEA	96.792	0.029	0.181
143	*G. morhua*	9	NS	53.904	0.382	0.414
144	*G. morhua*	8	NS	55.175	0.145	0.203
145	*G. morhua*	9	NEA	85.553	0.095	0.225
146	*G. morhua*	7	NS	94.861	0.14	0.035
147	*G. morhua*	8	NEA	97.864	0.053	0.257
148	*G. morhua*	9	NEA	82.832	0.162	0.33
149	*G. morhua*	9	NEA	58.421	0.098	0.126
150	*G. morhua*	9	NEA	94.691	0.038	0.19
151	*G. morhua*	9	NEA	96.711	0.117	0.429
152	*G. morhua*	9	NS	74.826	0.147	0.115
153	*G. morhua*	8	NS	98.809	0.344	0.061
154	*G. morhua*	8	NS	99.332	0.273	0.032
155	*G. morhua*	8	NS	90.052	0.287	0.171
156	*G. morhua*	9	NS	97.327	0.464	0.161
157	*G. morhua*	9	NS	74.569	0.49	0.359
158	*G. morhua*	9	NS	92.76	0.319	0.167
159	*G. morhua*	8	NS	96.99	0.164	0.044
160	*G. morhua*	9	NS	97.496	0.127	0.016
161	*G. morhua*	9	NEA	53.52	0.157	0.214
162	*G. morhua*	9	NS	90.839	0.266	0.154
163	*G. morhua*	9	NS	88.833	0.317	0.203

Supplementary Material – Peer reviewed paper references

The following papers are based on the work conducted during my PhD thesis:

Cusa M., Falcao L., De Jesus J., Biolatti C., Blondeel L., Bracken F.S.A., Devriese L., Garcés-Pastor S., Minoudi S., Gubili C., Acutis P.L., Mariani S. (2021) *Sustainability Science*. Fish out of water: Consumers' unfamiliarity with the appearance of commercial fish species.

Cusa M. *, St John Glew K. *, Trueman C., Mariani S., Buckley L., Neat F., Longo C. (2021) *Reviews in Fish Biology & Fisheries*. A future for seafood point-of-origin testing using DNA and stable isotope signatures. * *These authors contributed equally to the work*

Naaum A.M., **Cusa M.**, Hanner R.H., Helyar S., Elliott C., Mariani S., Wangh L.J., Sanchez A. (2021) *Food Research International*. FASTFISH-ID: A Rapid and Portable Method for Non-Targeted Authentication of Fish Species.

Supplementary Material – Ethics approval form

Research, Innovation and Academic Engagement Ethical Approval Panel

Doctoral & Research Support
Research and Knowledge Exchange,
Room 827, Maxwell Building
University of Salford
Manchester
M5 4WT

T +44(0)161 295 5278

www.salford.ac.uk/

7 June 2019

Marine Cusa

Dear Marine,

RE: ETHICS APPLICATION STR1819-46 – Advancing methods for genetic traceability of seafood in the European market

Based on the information you provided, I am pleased to inform you that your application STR1819-46 has been approved.

If there are any changes to the project and/ or its methodology, please inform the Panel as soon as possible by contacting S&T-ResearchEthics@salford.ac.uk

Yours sincerely,

Dr Devi Prasad Tumula
Deputy Chair of the Science & Technology Research Ethics Panel